The Theory of Determinants, Matrices, and Invariants

BLACKIE & SON LIMITED
66 Chandos Place, LONDON
17 Stanhope Street, GLASGOW

BLACKIE & SON (INDIA) LIMITED
103/5 Fort Street, BOMBAY

BLACKIE & SON (CANADA) LIMITED
TORONTO

The Theory of Determinants, Matrices, and Invariants

BY

H. W. TURNBULL, M.A., F.R.S.

Regius Professor of Mathematics in the United College, University of
St. Andrews. Formerly Scholar of Trinity College, Cambridge,
and Fereday Fellow of St. John's College, Oxford

SECOND EDITION

BLACKIE & SON LIMITED
LONDON AND GLASGOW

First issue 1928
Reprinted 1929
Second Edition, 1945
Reprinted 1946, 1948, 1950

Printed in Great Britain by Blackie & Son, Ltd., Glasgow

PREFACE

This book has grown out of a short series of lectures which were given in August, 1926, at the St. Andrews Congress of the Edinburgh Mathematical Society. It was the aim of those lectures to present in outline the salient features of the Invariant Theory, from its origins in the early forties of last century to the present day. But in the course of filling in the sketch, it was borne in upon me, more and more clearly, as the argument proceeded, that the subject takes its rise far earlier.

For this reason I have followed the method of Salmon in opening with an account of determinants. This also made it desirable to introduce the rudiments of another great department of algebra—the theory of matrices. These will chiefly be found in the first seven chapters, which have been written mainly with a view to their applications in what follows. It is no exaggeration to say that the well-known theorem, given by Laplace, for the development of a determinant, plays an essential part in all the main theorems of the symbolic invariant theory, as here adopted, with the one striking exception of the Basis theorem of Hilbert.

I am glad to acknowledge the great debt which mathematicians owe to Sir Thomas Muir for his charming *History of Determinants* which is at once a monument and an inspiration. If the present book encourages the reader to turn to the *History* and explore its farther fields, one of my objects will be attained. Here the subject is confined to what is called determinants in general and compound determinants. Perhaps the reader will also be tempted to dip into the buoyant papers of Sylvester (*Collected Works*) and the systematic treatise by Cullis (*Matrices and Determinoids*), who so generously displays the significance of the earlier writings by Sylvester.

As to the invariant theory itself, an attractive approach to binary and ternary forms has for many years been accessible through the admirable treatises by Elliott (*The Algebra of Quantics*, Oxford, second edition, 1908) and by Grace and Young (*The Algebra of Invariants*, Cambridge, 1902), the former developing the direct, and the latter the symbolic methods. But during the present century considerable advances have been made in studying quaternary and higher forms (involving four or more homogeneous variables), both in the algebra itself and in its application to physics through the concept of Relativity. Accordingly, while I have incorporated just enough of the binary theory to give a short connected exposition of its developments, my chief concern has been with the general forms.

Whatever completeness may attach to the present argument is finally due to the memoirs and recent books by Weitzenböck[1] and Study.[2] To the former belongs the credit of extending the work of Clebsch and Gordan from the binary to the general case. But perhaps the most remarkable service which he has hitherto rendered is to give a complete account of the basis of analytical projective geometry in relation to all the usual metrical forms, Euclidean and others. An exposition of these results is given in Chapter XXI.

In such a far-flung theory, with all its great ramifications into pure algebra, the theory of groups, projective and differential geometry, somewhere or other the line must be drawn: and this has been done as follows. First, beyond a bare introduction to each (Chapters XX and XXI), the two chief applications, to algebraic and differential geometry, have been omitted. However logically appropriate fuller treatment would have been, it was felt that justice could not be done to what is an extraordinarily attractive and penetrating type of analytical geometry in three-fold and higher space, at the end of a long algebraic theory. But the reader can find a full account of the plane geometry in the later chapters by Grace and Young.

Secondly, there is no mention of the interesting algebra of *alternate* numbers for which $a \times b = - b \times a$. These have a long historical record in the work of Grassmann, Whitehead, Scott and Matthews, and others. The omission calls for some explanation, because in the deft hands of Dr. Weitzenböck, a key to the general invariant theory is provided by *complex symbols*, which are a type of alternate numbers. But it was found that by enlisting the full implications of Sylvester's Theorem (1851) (p. 48), the ordinary symbols provide quite a natural medium for the whole general theory, from beginning to end.

---

[1] *Invariantentheorie* (Groningen, 1923).

[2] *Einleitung in die Theorie der Invarianten linearer Transformationen auf und der Vectorenrechnung* (Braunschweig, 1923).

There is also no attempt to grapple with all the details in the theory of canonical forms and invariant factors; but the necessary suggestions for further reading have been made at suitable stages. Neither has room been found for the discussion of special complete systems; nor for the extensive theory of modular invariants which have lately received great attention in America.

Here and there, illustrative examples have been included, often as straightforward applications but occasionally as more advanced problems and suggestions for further inquiry and research. Among examples of determinants and matrices are several for which I am indebted to Professor E. T. Whittaker and Dr. A. C. Aitken.

My best thanks are due to my colleague, Dr. W. Saddler, for his ripe judgment and criticism in reading the work, and for offering many valuable suggestions; and to Dr. J. Williamson for reading the proof-sheets and giving further helpful advice; and also to Dr. J. Dougall for his expert and very efficient help in removing both mathematical and typographical blemishes.

H. W. TURNBULL.

St. Andrews, June 1928

PREFACE TO THE SECOND EDITION

Besides a revision of the existing work where numerous corrections have been made, two chapters are added. Chapter XXII contains many notes upon determinants and matrices which experience has shown to be useful and appropriate to invariant theory. Bideterminants provide a new algorithm for investigating problems both elementary and advanced. Chapter XXIII gives a brief survey of Young's theory of standard forms, and thus continues a theme hinted at in Chapter XXI. It is designed to point the way to further investigation into the significance of the Gordan-Capelli series for the later and abstract concepts initiated by Frobenius and Schur. References are given at the close of the chapter to the important developments by H. Weyl, D. E. Littlewood, D. E. Rutherford, and others, either in the abstract or in the algebraical or in the geometrical theory.

I take this opportunity of mentioning an interesting series of papers which Weitzenböck and Bos continue to communicate to the Amsterdam Academy, upon the differential geometry and invariants of a ruled surface in four dimensions, thus providing a fruitful example of invariants in quinary forms.

My best thanks are again due to the many friends who have called my attention to blemishes in the earlier edition, and particularly to Dr. A. C. Aitken and to Dr. J. Dougall and the publishers for helping with their mathematical experience and technical skill.

H. W. TURNBULL.

St. Andrews, January, 1945.

CONTENTS

CHAPTER I

MATRICES AND DETERMINANTS

Page
1. Notation - - - - - - - - - - 1
2. Definition of Matrix - - - - - - - - 2
3. The Transposed Matrix - - - - - - - - 5
4. System of Linear Equations - - - - - - - 6
5. Linear Combinations of Rows or Columns. Number Field. Rank 8
6. Linear Equations which are not Homogeneous - - - 10
7. Condition of Solubility - - - - - - - - 11
8. Number Rings - - - - - - - - - 12

CHAPTER II

FUNDAMENTAL PROPERTIES OF THE DETERMINANT

1. Derangements - - - - - - - - - 13
2. The C_+ and C_- Classes - - - - - - - - 15
3. Definition of Determinant - - - - - - - 17
4. Arrangement of Terms in the Expansion of a Determinant. Standard Order Co-factors - - - - - - 19
5. Laplace's Development of a Determinant - - - - 22
6. Algebraic Complements and Minors of Order r - - - - - 26
7. Determinantal Permutation - - - - - - - 27

CHAPTER III

LINEAR PROPERTIES. FUNDAMENTAL LAPLACE IDENTITIES

1. Linearity. Homogeneity - - - - - - - 30
2. Special Determinants - - - - - - - - 32
3. Double Suffix Notation and other Contractions - - - 32
4. A Determinant is irresoluble into Factors - - - - 33

Page
5. Rules for Combining Matrices - - 34
6. Currency of a Matrix - - - - - 37
7. Transposition Properties of Determinants 38
8. Fundamental Laplace Identities - - 41
9. Fundamental Identities of Order n - - 44
10. Implicit and Explicit Convolution - - 46
11. General Fundamental Identities of Order n 47
12. Linear Relation between $n + 1$ Linear Forms 50
13. Principle of Duality - - - - 51

Chapter IV

MULTIPLICATION OF MATRICES AND DETERMINANTS

1. Fundamental Laws of Algebra - - - - - 57
2. The Law of Multiplication of Matrices - - - - 59
3. Product of Square Matrices of Order n - - - - 61
4. Double Suffix Notation of Multiplication - - - 63
5. The Division Law - - - - - - - 64
6. Products of Determinants - - - - - - 65
7. Reciprocal and Adjugate Determinants - - - - 66
8. The Index Law and the Reversal Law of a Matrix - - 68
9. Summary of Laws of Matrices - - - - - 70

Chapter V

LINEAR EQUATIONS. THE THEOREM OF CORRESPONDING MATRICES. FURTHER THEOREMS

1. Matrices and Linear Equations. Rank - - - - - - 73
2. Application to Linear Equations - - - - - - 75
3. The Upper Suffix Notation - - - - - - - - 77
4. The Theorem of Corresponding Matrices (The Binet-Cauchy Theorem) - - - - - - - - - - 79
5. Inner Product of Two Rectangular Matrices. Bideterminant - 82
6. Laplace Developments of the Inner Products - - - - - 83
7. Rank of the Product of Matrices - - - - - - 84
8. The Simplex - - - - - - - - - - 84
9. Extended Form of Cauchy's Theorem, commonly called Sylvester's Theorem on Compound Determinants - - - 87
10. The Generalized Ratio Theorem - - - - - - 89
11. Tensor Constants of the Fundamental Identities - - - 90
12. Application of the Principle of Duality - - - - - 92
13. The Sylvester Identity - - - - - - - - 93
14. Formal Proof of the Sylvester Identity - - - - - 95

CHAPTER VI

SPECIAL TYPES OF DETERMINANT

Page

1. Properties of Matrices and Determinants connected with the Leading Diagonal - - - - - - - - 98
2. The Cayley Hamilton Theorem - - - - - - 99
3. Special Types of Determinant - - - - - - 101
4. Reciprocation of Bordered Determinants - - - - 102
5. Bordered Adjugate Determinant - - - - - - 104
6. Symmetrical Matrices and Determinants - - - - - 104
7. Skew Symmetric Determinants - - - - - - 105
8. Characteristic Function of a Skew Matrix - - - - 107
9. Summary of Theorems on Compound Determinants - - - 107

CHAPTER VII

DIFFERENTIATION OF A DETERMINANT

1. The Polarizing Process - - - - - - - - - 110
2. The Capelli Operators - - - - - - - - - 112
3. The Cayley Operator - - - - - - - - - 114
4. Theorem of Corresponding Matrices adapted to the Capelli Operator - - - - - - - - - 116
5. Connexion between Substitutional Analysis and Differentiation 119
6. Jacobians - - - - - - - - - - 124
7. Rank of Jacobian Matrix - - - - - - - - 126

CHAPTER VIII

BINARY FORMS

1. Binary Invariants - - - - - - - - - 128
2. Orthogonal Transformation and Invariants - - - - 130
3. Development of the Invariant Theory - - - - - 132
4. The Binary Form or Quantic - - - - - - - - 133
5. Gradient. Degree and Weight - - - - - - - 134
6. The Induced Linear Transformation of the Binary n-ic - - 135
7. Polar Forms - - - - - - - - - - 137
8. Formal Definition of Invariant - - - - - - - 138
9. Simultaneous Invariants - - - - - - - - 139
10. The Aronhold Operator - - - - - - - - 140
11. Multilinear Invariants - - - - - - - - - 142
12. Covariants - - - - - - - - - - 143
13. Relation between Linear Forms and Covariants - - - 145

CHAPTER IX

THE GENERAL LINEAR TRANSFORMATION

Page
1. Cogredience and Contragredience - - - - - - 147
2. Linear Transformations in Matrix Notation - - - - 149
3. Orthogonal Transformations and Matrices - - - - 152
4. Cayley's Determination of the Orthogonal Matrix all whose Latent Roots differ from Negative Unity - - - - 155
5. Linear Transformation with Absolute Quadric - - - - 158
6. Group of the Orthogonal Matrix - - - - - - 160
7. Dimensions of the Transformation Group - - - - 161
8. Induced Compound Transformations - - - - - 163
9. Connexion between Matrices and Quaternions - - - - 166

CHAPTER X

GENERAL PROPERTIES OF INVARIANTS

1. Linear Transformation of the General Form of Order p - - 168
2. Projective Invariants - - - - - - - - 169
3. Homogeneity of Invariants - - - - - - - 171
4. Ground Forms - - - - - - - - - 172
5. Symbolic Notation - - - - - - - - 173
6. Symbols for Forms in Three or More Variables - - - 175
7. Polar Forms - - - - - - - - - 177
8. Equivalent Symbols - - - - - - - - 179

CHAPTER XI

THE FIRST FUNDAMENTAL THEOREM

1. Symbolic Factors. Inner and Outer Products - - - 182
2. Effect of Linear Transformation on the Symbols - - - 183
3. Converse Theorem - - - - - - - - 184
4. The Valency Condition $pq = nw + \varpi$ for Single Ground Form - 186
5. First Fundamental Theorem for a System of Linear Forms - 187
6. Invariants of One or More General Ground Forms - - - 189
7. Examples of Invariants. Interchange of Equivalent Symbols - 191
8. Double Convolution of Symbols referring to a Quadric Form - 193
9. Solution of Symbolic Linear Equations - - - - - 195

CHAPTER XII

MULTILINEAR FORMS

Page
1. Multilinear Forms - - - - - - - - 197
2. Symbolic Representation of Multilinear Forms - - - 198
3. Classification of Multilinear Forms - - - - - 199
4. Cogredient and Contragredient Symbols - - - - - - 200
5. Equivalent Symbols - - - - - - - - - 201
6. Effect of Linear Transformation on the Symbols - - - 201
7. Fundamental Theorem for the General Multilinear Form - 203
8. Covariants, Contravariants, and Mixed Concomitants - - 206
9. Convolution and Resolution - - - - - - - - 207
10. The Fundamental Theorem for the General Case - - - 208
11. Proof of the Fundamental Theorem - - - - - - 210

CHAPTER XIII

SYMBOLIC METHODS OF REDUCTION

1. The Fundamental Identities - - - - - - - - 213
2. The Second Fundamental Theorem - - - - - - 214
3. Binary Quadratic Forms. Reducibility - - - - - - 215
4. Significance of the Complete System - - - - - - 218
5. Canonical Form of Two Binary Quadratics - - - - 219
6. Extension to Forms of Higher Order - - - - - 221
7. Transvectants - - - - - - - - - - 221
8. Reducibility of Jacobians - - - - - - - - 223
9. Remarks on the Proof of the Second Fundamental Theorem - 225

CHAPTER XIV

SEMINVARIANTS. ALGEBRAICALLY COMPLETE SYSTEMS

1. Seminvariants and Leading Term of a Concomitant - - 226
2. Seminvariants as Solutions of Partial Differential Equations - 227
3. Algebraically Complete Systems. Syzygies - - - - 231
4. Irreducibility. Gordan's Theorem - - - - - - - 233

CHAPTER XV

THE GORDAN-HILBERT FINITENESS THEOREM

1. Hilbert's Basis Theorem - - - - - - - - 235
2. Proof of Gordan's Theorem - - - - - - - - 238
3. Limit to the Number of Syzygies - - - - - - - 239

Page
4. Multiple Fields - - - - - - - - 240
5. Combinants - - - - - - - - - 242
6. Further Examples of Complete Systems. The Binary Cubic 244
7. The Binary Quartic Form - - - - - - 245
8. References to Complete Systems - - - - - 246

Chapter XVI

CLEBSCH'S THEOREM

1. Introduction of Clebsch's Theorem - - - - 248
2. Compound Polars. Standard Forms - - - 249
3. Reduction to Standard Form - - - - 250
4. The Gordan-Capelli Series - - - - - 253
5. Examples of the Series for Binary and Ternary Fields 255
6. Normal Forms - - - - - - - 255
7. Historical Note - - - - - - - 258

Chapter XVII

APPLICATIONS OF CLEBSCH'S THEOREM. APOLARITY AND CANONICAL FORMS

1. Similar Forms - - - - - - - 259
2. Types - - - - - - - - 260
3. Peano's Theorem - - - - - - - 261
4. Dual Similar Forms - - - - - - 262
5. Apolarity - - - - - - - - 264
6. Apolarity of Dissimilar Forms - - - - 264
7. Canonical Forms - - - - - - - 265
8. Counting Constants is not Sufficient - - - 267
9. Proof of the Lasker-Wakeford Theorem - - - 268

Chapter XVIII

INVARIANT EQUATIONS AND GRAM'S THEOREM

1. Expression of a Gradient by Coefficients of Covariants - 270
2. Invariant Equations - - - - - - - 271
3. Gram's Theorem - - - - - - - 271
4. Grace's Theorem - - - - - - - - 273
5. Invariants as Elimination Results - - - - 274
6. The Equivalence Problem - - - - - - 277
7. Extension of Stroh's Lemma - - - - - - 278

CHAPTER XIX

GEOMETRICAL INTERPRETATIONS OF ALGEBRAIC FORMS

Page
1. Homogeneity and Correspondence - - - - - - - 280
2. Principle of Duality - - - - - - - - - 282
3. Further Binary Results - - - - - - - - 284
4. Connexion of Binary with Higher Fields - - - - 286
5. The Clebsch Transference Principle. Extensionals - - - 287
6. Projective Properties - - - - - - - - - 290
7. First Geometrical Interpretation of Linear Transformation. Collineation - - - - - - - - - - 291
8. Latent Points of a Transformation - - - - - - - 292
9. Second Geometrical Interpretation of Linear Transformation. Change of Frame of Reference - - - - - 293
10. Reciprocation and Correlation - - - - - - - 294
11. Correlation - - - - - - - - - - - 295
12. Canonical Form of a Matrix - - - - - - - - 295

CHAPTER XX

THE GENERAL QUADRIC

1. Complete System of the General Quadric - - - - - 297
2. Self-conjugate Simplex - - - - - - - - - 299
3. Canonical Form of the Quadric - - - - - - - 300
4. Theory of Two Quadrics - - - - - - - - 301
5. Reduction of Two Quadrics to Sums of Squares - - - 302
6. Complete System of $(n + 1)$ Invariants - - - - - 304
7. Complete Systems involving Variables - - - - - 306

CHAPTER XXI

MISCELLANEOUS RECENT DEVELOPMENTS

1. Restricted Transformations - - - - - - - - 309
2. Preparatory Reductions leading to the Proof of the Fundamental Theorem - - - - - - - - - - 310
3. Characteristic Invariant Property - - - - - - - 312
4. Proof of the First Fundamental Theorem - - - - 312
5. Consequences of the Theorem - - - - - - - 314
6. The Orthogonal Group - - - - - - - - - 315
7. Fundamental Theorem of Orthogonal Transformation - - 316
8. First Fundamental Theorem for Proper Orthogonal Invariants - 318

Page
9. The Hermitian Transformation with an Absolute Quadric - 320
10. Geometrical Significance of the Adjunction Theorem - 323
11. Remarks on the Adjunction Theorem - - - - 325
12. Connexion between Differential and Projective Invariants - 326
13. Prepared Systems - - - - - - - - 329
14. Quantitative Substitutional Analysis - - - - 330

CHAPTER XXII

MISCELLANEOUS PROPERTIES OF DETERMINANTS

1. Remark on the Fundamental Identities - - - - 334
2. Partitions by Rows and Columns. Double Currency - - 335
3. Conformable Partitions for Addition and for Multiplication - 335
4. Schweins' Theorem - - - - - - - - 336
5. Further Use of the Unit Matrix - - - - - - 338
6. Schweins' Series - - - - - - - - - - 339
7. Dodgson's Method of Evaluating a Determinant by Condensation - - - - - - - - - - 340
8. The Bordered Unit Matrix and the Binet-Cauchy Theorem - 341
9. Bideterminants - - - - - - - - - - 342
10. Note on Matrices - - - - - - - - - - 346
11. The Case of Failure of Cayley's Orthogonal Matrix - - - 348
12. Matrix Representation of the Binary Symbols - - - - 349

CHAPTER XXIII

STANDARD FORMS

1. Compound Polarization - - - - 353
2. Young's Standard Forms and Tableaux - - 355
3. Young's Standard Theorem - - - - 357
4. Enumeration of Standard Forms - - - 358
5. Double Standard Forms - - - - 360
6. Polarized Double Standard Forms - - - 362

INDEX - - - - - - - 365

The Theory of Determinants, Matrices, and Invariants

CHAPTER I

MATRICES AND DETERMINANTS

1. Notation.

The fundamental importance of determinants as working tools in mathematics has come to be so widely recognized that it may be assumed that the reader has some practical knowledge of them, and in particular that he has realized their value in providing a simple general rule for the solution of linear equations. Certain introductory results may therefore be given without undue emphasis on intermediate steps, which can easily be supplied. Our first object is to learn a notation and a few important definitions.

Suppose there are two homogeneous linear equations in three variables x, y, z,

$$\left.\begin{aligned} a_1x+b_1y+c_1z&=0,\\ a_2x+b_2y+c_2z&=0.\end{aligned}\right\} \quad . \quad . \quad . \quad . \quad (1)$$

Then in general they have a solution

$$\frac{x}{b_1c_2-b_2c_1}=\frac{y}{c_1a_2-c_2a_1}=\frac{}{a_1b_2-a_2b_1}. \qquad (2)$$

These denominators, which are called *determinants of the second order*, can be written shortly in various ways, all of which have great value:

$$\left.\begin{array}{llll} \text{(i)} & |\,b_1c_2\,|\,, & |\,c_1a_2\,|\,, & |\,a_1b_2\,|\,,\\ \text{(ii)} & (bc)_{12}\,, & (ca)_{12}\,, & (ab)_{12}\,,\\ \text{(iii)} & (bc)\,, & (ca)\,, & (ab)\,. \end{array}\right\} \qquad (3)$$

The last of these ways makes use of the obvious fact that if two letters bc are written down side by side, one is first and the other is second, read from left to right. We agree to drop the suffixes in (iii), whenever they are 1, 2, for exactly the reason that we drop the index 1 in writing a^p when $p = 1$. In fact we define $(bc)_{ij}$ to mean $b_i c_j - b_j c_i$ and merely suppress the suffixes ij in the case when $i = 1$, $j = 2$.

A fourth and more familiar notation for the determinant $b_1 c_2 - b_2 c_1$ is the well-known square array, introduced by Cayley[1] in 1841 long after determinants (and much that will concern us in this chapter) were first invented. It is

$$\begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} \tag{4}$$

which has the advantage of showing such coefficients of the original equations, as appear in the first determinant, exactly in their same relative positions.

This leads to still more ways, all useful, of writing down the solution of equations (1):—

$$\left.\begin{aligned} &\text{(i)}\ \ x : y : z = \begin{vmatrix} b_1 & c_1 \\ b_2 & c_2 \end{vmatrix} : \begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix} : \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} . \\ &\text{(ii)}\ \ x : y : z = (bc) : (ca) : (ab) , \\ &\text{(iii)}\ \ x,\ y,\ z \propto \left\| \begin{matrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{matrix} \right\| . \end{aligned}\right\} \tag{5}$$

In each of these cases three equations have been grouped into one statement. Only in (iii) we note that an essentially new idea is present: the double vertical lines,[2] before and after the rectangular array, signify that determinants are to be chosen therefrom by suppressing in turn the first, second, and third column of letters, and at the same time retaining the orders b, c; c, a; a, b of the columns.

2. Definition of Matrix.

There is obvious importance in adopting a methodical arrangement of equations and all such polynomial expressions, involving several variables x, y, z. Also, because of the convenient fact that many of the properties of a square or oblong formation can be illustrated by arranging four or six things two by two in a square, or two by three in an oblong, we can continue to extract useful general notions from our equations (1). The set of coefficients

---

[1] In 1841, *Collected Works*, **1**, 1.

[2] This notation has sometimes also been used to denote the matrix of the array.

$$\begin{matrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{matrix}$$

of (1), arranged in their relative positions, is an example of a matrix of orders two and three. *A matrix of orders* m *and* n *simply means a set of* mn *numbers arranged in rectangular array with* m *rows and* n *columns.*

At first sight such a definition strikes one as awkward and vague, for the question naturally arises in the mind, what shall we *do* to these numbers, shall we add or subtract them or form them into determinants? Nevertheless it is exceedingly useful to train ourselves to think of an array of numbers as a single thing with properties of its own, and to hold ourselves in readiness to operate on the terms or elements of the array in any convenient way that suggests itself, as in fact we have done in the preceding results (2), (3), (5). We are indeed all familiar with this idea, for ordinary Cartesian co-ordinates

$$[x, y, z]$$

of a point in space provide a simple instance. Here the matrix is of orders one and three. This involves more than merely three numbers x, y, z; it is three numbers together with a specific relation between them; namely, that they are ordinally arranged. In general when x, y, z differ, the geometrical interpretation of the different arrangements

$$[x, y, z],\ [x, z, y],\ [y, x, z],\ [y, z, x],\ [z, x, y],\ [z, y, x]$$

is six different points: and this is hint enough that regarded as algebraic elements (molecules, if we like), we may with advantage study the behaviour of matrices, always treating them as single integral things, and not as elaborate clusters of component parts. Just as Cayley first provided us with the well-known square notation for determinants (Cf. (4)) so also we have to thank him for first[1] enunciating this principle. He, however, confined the definition of a matrix to a square formation only.

Let us agree to use brackets [ ] for enclosing the constituents of a matrix, and incidentally for expressing co-ordinates of a point, in plane or space, so that we can now proceed to discuss the matrix M of the coefficients of linear equations (1), and write it

$$M = \begin{bmatrix} a_1, & b_1, & c_1 \\ a_2, & b_2, & c_2 \end{bmatrix}.$$

We can also with advantage notice that there is a matrix X of the homogeneous variables x, y, z in the equations, namely

$$X = [x, y, z].$$

It is a simple but far-reaching fact that for a given system of equations, arranged by columns and rows as in (1), there are these two matrices M and X. One cannot exist without the other.

It was said that *in general* equations (1) have a solution. By this is meant all cases in which the two equations are effectively distinct, a state of things that only breaks down if

$$a_1 : b_1 : c_1 = a_2 : b_2 : c_2.$$

When this happens the coefficients of one equation are proportional to those of the other, and the two equations furnish no more information about x, y, z than either of them alone would do. It is then impossible to derive solutions (2) from (1), still less the results (5). If we define the phrase *determinants of the matrix* M to mean all the determinants (bc), (ca), (ab), we may state that *the equations* (1) *are soluble unless all the determinants of the coefficient matrix* M *are zero.*

Suppose *two* of the determinants (bc), (ca) vanish. Then, eliminating c_1, c_2 it follows that (ab) also vanishes. Hence a sufficient condition, for the insolubility of equations (1) in the form (2), is that two of the three determinants of M vanish.

If, however, only *one* determinant vanishes, (bc) say, then $x = 0$, $y : z = (ca) : (ab)$. And we may define equations (2) to have this meaning, although standing alone the ratio $x : (bc)$ would now be meaningless and could not be employed.

Just as there is geometrical significance in [0, 0, 0] which denotes the set of co-ordinates of the origin, so we may presume that the *null matrix*

---

[1] *Phil. Trans.* (1858); *Collected Works*, **2**, 475.

$$O = \begin{bmatrix} 0, & 0, & 0 \\ 0, & 0, & 0 \end{bmatrix}$$

has algebraic significance, although in relation to equations (1) it appears to indicate their non-existence.

We may sum up this little investigation by attaching a special term *rank* to a matrix. *The rank of*

$$\begin{bmatrix} a_1, & b_1, & c_1 \\ a_2, & b_2, & c_2 \end{bmatrix}$$

is two, unless all the determinants $b_1c_2 - b_2c_1$, $c_1a_2 - c_2a_1$, $a_1b_2 - a_2b_1$ vanish, in which case it is *one,* unless again all six elements a_1, b_1, c_1, a_2, b_2, c_2 vanish, in which case it is *zero.*

EXAMPLES

1. If $a_1x + b_1y + c_1z = 0$, $a_2x + b_2y + c_2z = 0$ are the Cartesian equations of two distinct planes, prove the rank of the coefficient matrix is two. If the rank is one, what is known about the planes?

2. If these equations refer to lines in a plane, in areal (or other homogeneous) co-ordinates, what is the significance of the rank of their matrix?

3. A two-by-three matrix has rank unity. Show that if its rows denote areal co-ordinates of a point, each row denotes the same point.

What other two-by-three matrix has rank unity?

Ans. A matrix in which one row is three zeros.

4. A three-by-two matrix has three rows and two columns. If each row is interpreted as Cartesian co-ordinates of a point in a plane, show that its three determinants all vanish if the three points are in line with the origin.

3. The Transposed Matrix.

If we interchange columns and rows without disturbing the order of either, reading columns downwards and rows from left to right, we obtain the *transposed* [1] matrix. Let us use an accent to denote the transposed matrix. Thus the transposed of M is

$$M' = \begin{bmatrix} a_1, & a_2 \\ b_1, & b_2 \\ c_1, & c_2 \end{bmatrix}$$

---

[1] Sometimes called *conjugate* matrix.

If we transpose M' we obtain M, so that here we have an example of a *conjugate* or *symmetrical* relation between two things. So also the transposed of $X = [x,\ y,\ z]$ is

$$X' =$$

and vice versa. The determinants of the matrix M' are (bc), (ca), (ab) which are the same as those of M. More precisely the relation

$$\begin{matrix} b_1 & b_2 \\ c_1 & c_2 \end{matrix} \qquad \begin{matrix} b_1 & c_1 \\ b_2 & c_2 \end{matrix}$$

shows the identity of corresponding transposed determinants. But we should naturally think of the determinants of M as forming a *row* (i.e. a matrix) of three elements, of the same pattern as X,

$$[\,(bc),\ (ca),\ (ab)\,]$$

while those of M' form a *column* of three.

Owing to the practice of writing from left to right, rather than, as the Chinese do, from top to bottom, we have never accustomed ourselves to thinking of co-ordinates of a point written downwards as in X'. It will later appear that this novel way sometimes has very great advantages. But occasionally, in order to save space, a column matrix will be written horizontally and enclosed in brackets $\{\ \ \}$. Thus

$$X' = \{x,\ y,\ z\}.$$

4. System of Linear Equations.

Before dealing with the general case involving n variables, let us consider a set, or system, of three linear equations homogeneous in four variables

$$\begin{aligned} a_1 x + b_1 y + c_1 z + d_1 t &= 0, \\ a_2 x + b_2 y + c_2 z + d_2 t &= 0, \\ a_3 x + b_3 y + c_3 z + d_3 t &= 0. \end{aligned} \quad . \quad . \quad . \quad (6)$$

Multiplying these respectively by $(bc)_{23}$, $(bc)_{31}$, $(bc)_{12}$ and adding we find that all terms involving y or z disappear, and that the result may be written

$$(abc)\,x + (dbc)\,t = 0, \quad . \quad . \quad . \quad . \quad . \quad (7)$$

where (abc) is a convenient symbol for the coefficient of x in the result, and (dbc) for that of t. Thus

$$\begin{aligned}(abc) &= a_1(bc)_{23} + a_2(bc)_{31} + a_3(bc)_{12}\\ &= a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1. \quad (8)\end{aligned}$$

Likewise

$$(dbc) = d_1b_2c_3 - d_1b_3c_2 + d_2b_3c_1 - d_2b_1c_3 + d_3b_1c_2 - d_3b_2c_1.$$

Manifestly the series for (abc) may also be written

$$a_1b_2c_3 - a_1c_2b_3 + b_1c_2a_3 - b_1a_2c_3 + c_1a_2b_3 - c_1b_2a_3, \qquad (9)$$

and further, it is clear, on expansion, that the following equalities are true

$$(abc) = (bca) = (cab) = -(acb) = -(bac) = -(cba). \qquad (10)$$

Before Cayley introduced the notation

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} \qquad (11)$$

for this series (8), which is a *determinant of the third order*, it was frequently written

$$\Sigma \pm a_1b_2c_3, \quad . \; . \; . \; . \; . \; . \; . \quad (12)$$

the summation indicating *either* that a, b, c are to be deranged in all six possible ways, as in (9), without deranging the suffix order 1, 2, 3, *or* vice versa, as in series (8), the suffixes are deranged but the letters are not. The ± sign here indicates that some terms have a positive sign and some a negative, the choice depending on a rule to be presently explained.

If we also solve for y or z, as in (7), we obtain in general

$$\frac{x}{(bcd)} \qquad (acd) - (abd) \qquad \frac{-t}{(abc)}, \qquad (13)$$

as should be carefully verified. The negative signs occurring with the alternate variables y, t are inserted to maintain the alphabetical order in the denominators. For these determinants are obtained in the Cayley notation by suppressing in turn each one of the columns of the coefficient matrix

$$\begin{bmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{bmatrix} \tag{14}$$

Some writers define these as the determinants of this matrix: it is preferable, however, to attach a sign + or — according as an odd or even column is suppressed. Thus x, y, z, t *are respectively proportional to the determinants* A, B, C, D *of the coefficient matrix, namely*

$$A = (bcd), \quad B = -(acd), \quad C = (abd), \quad D = -(abc). \tag{15}$$

5. Linear Combinations of Rows or Columns. Number Field. Rank.

It is convenient to have a precise notation applicable to matrices, determinants, and systems of equations. A few examples suffice to explain it. Consider the array

$$\begin{array}{ccc} a & b & c \\ x & y & z \\ a+x & b+y & c+z. \end{array} \tag{16}$$

Here we obtain the third row by adding the elements of the two other rows columnwise. This is denoted by

$$\text{row}_3 = \text{row}_1 + \text{row}_2. \tag{17}$$

Again consider the array of four columns,

$$\begin{array}{cccc} a & b & c & a+b+c \\ x & y & z & x+y+z. \end{array} \tag{18}$$

Here is an example of adding row-wise. We denote it by

$$\text{col}_1 + \text{col}_2 + \text{col}_3 = \text{col}_4. \quad . \quad . \quad . \tag{19}$$

Next the two arrays

$$\begin{array}{ccc} a & b & c \\ pa & pb & pc \end{array} \qquad\qquad \begin{array}{cc} a & qa \\ b & qb \\ c & qc \end{array}$$

exhibit what is meant by multiplying a row or column by a given number. We write these

$$\text{row}_2 = p\ \text{row}_1 \qquad \text{col}_2 = q\ \text{col}_1$$

respectively.

In general, by

$$p\ \text{row}_1 + q\ \text{row}_2 + r\ \text{row}_3 \quad . \quad . \quad . \quad (20)$$

is meant: *form a new row by multiplying the first by* p, *the second by* q, *the third by* r, *and adding.* Similar remarks apply to columns, but only the former process applies immediately to equations, as for instance in reaching result (7) of the last section from (6).

We shall assume that all the symbols $a, b, c, x, p \ldots$, hitherto used, stand for real or complex numbers. It follows that the process (20) implicitly includes subtraction as well as addition of rows, since one or other of p, q, r may be real and negative.

At this stage it is useful to have a clear conception of what is meant by a *field* of numbers. This can be defined as follows.

Definition of Number Field.—*A class of* two or more *complex numbers forms a field if whatever two equal or unequal members* p *and* q *are chosen then* p + q, p − q, p × q, p ÷ q *are themselves members of the field, excepting the case* q = 0 *in the quotient* p ÷ q.

It follows that integers do not form a field, because for instance $2 \div 3$ is excluded: but rational numbers form a field. So also do real numbers. So also do numbers of the type $a + b\sqrt{5}$ where a and b are rational. So also do complex numbers. It also follows that zero is a member of every possible field, by taking p and q equal in $p - q$.

Suppose we now prescribe a definite field F for the numbers p, q, r of (20). By this procedure we are said to form *a linear combination of the rows or columns* in question.

Obviously if p, q, r are all zero we should form a row of zeros from

$$p\ \text{row}_1 + q\ \text{row}_2 + r\ \text{row}_3.$$

But, excluding this case, suppose we still get a row of zeros, i.e. *a null row,* when not all p, q, r vanish; then we term the several rows so combined *linearly related* or *linearly dependent in the* *field F.* The same definition applies to columns. And if we cannot get a null row unless all p, q, r vanish, the several rows so combined are *linearly independent in the field F.* Likewise for any number of rows, and columns.

This distinction between linear dependence and independence is of the utmost importance, and should be carefully thought out with these simple cases, in order to pave the way for its more elaborate use at later stages.

We can now utilize this discussion of linearity to define the *rank* of a matrix in general. The rank of a square matrix is the greatest number of its rows or columns which can be found to be linearly independent. That of a rectangular matrix with fewer rows than columns is the greatest number of its rows which are linearly independent. If there are more rows than columns, the same test applies to its columns.

So for the $m \times n$ matrix the rank may be any whole number $0, 1, 2, \ldots, r$ not exceeding either m or n. It will be shown in Chapter V that this definition amounts to the same thing as that already adopted in §2.

6. Linear Equations which are not Homogeneous.

The solutions of the homogeneous equations in either three or four variables already treated give the ratios but not the exact values of the variables $x, y, z, \ldots$. In general n such equations for $n + 1$ unknowns $x, y, z, \ldots, t$ determine the ratios

$$x : y : z : \ldots : t$$

in terms of the coefficients. It follows that if one of these, the last, t, for example, is given in value, the rest can be determined. It is useful to take $t = -1$, for when this is done we can at once write down solutions as follows. The *binary* case has two equations for two unknowns x, y : the *ternary* case has three for three unknowns.

(1) *The Binary Case.*

If
$$\left.\begin{aligned} a_1 x + b_1 y &= c_1 \\ a_2 x + b_2 y &= c_2 \end{aligned}\right\} \qquad (ab) \neq 0,$$

then
$$\frac{x}{(cb)} = \frac{y}{(ac)} = \frac{1}{(ab)} \quad \text{so that} \quad x = \frac{(cb)}{(ab)}, \quad y = \frac{(ac)}{(ab)}.$$

(2) *The Ternary Case.*

If

$$\left.\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}\right\} \qquad (abc) \neq 0,$$

then

$$\frac{x}{(dbc)} = \frac{y}{(adc)} = \frac{z}{(abd)} = \frac{1}{(abc)},$$

so that

$$x = \frac{(dbc)}{(abc)}, \quad y = \frac{(adc)}{(abc)}, \quad z = \frac{(abd)}{(abc)}.$$

It is worth while noticing the simple manner in which these last fractions, giving x, y, z, are formed. Each denominator is the determinant formed from the left-hand side coefficient matrix of the given system of equations. This matrix is now a *square* array. The numerator of x is obtained from the same determinant by suppressing the first column and substituting the column of d's in its stead. By substituting in the second and third columns of (abc) similarly, we obtain the respective numerators of y and z. This device overcomes the difficulty (cf. (14)) of affixing the sign. It has the advantage of perfect generality, for it applies equally well to n equations involving determinants of the nth order. This rule was first given by Cramer in 1750.

7. **Condition of Solubility.**

These equations are soluble, as we see, unless in the binary case $(ab) = 0$, and in the ternary case $(abc) = 0$. A similar test holds for more variables. Thus the equations in the ternary case are soluble if the rank of the square matrix

$$\begin{array}{ccc} a_1 & b_1 & c_1 \\ & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{array}\rfloor$$

is three. It is interesting to examine the cases of failure, when the rank is less than three; but as our chief concern is with the soluble case we leave this aside.

8. **Number Rings.**

It is useful to give a name to classes of numbers, such as the real integers, or again the complex integers, which fail to form a field only by disobeying the division law. Such a class is called a *ring* and can be defined as follows.

Definition of a Number Ring. A class of two or more numbers forms a ring if whatever two equal or unequal members p and q are chosen then $p + q$, $p - q$, $p \times q$ are themselves members of the ring.

It follows that integers form a ring but not a field. The idea of the ring has place both in elementary arithmetic and in the theory of numbers. Nor is division entirely excluded, if it arises out of repeated subtraction, and involves a remainder (within the ring) which may happen to be zero.

EXAMPLES

1. Write down the determinants of the following arrays:

$$\begin{bmatrix} 1, & 2, & 3 \\ 4, & 5, & 6 \end{bmatrix}, \quad \begin{bmatrix} a_2 & a_3 & a_4 \\ b_2 & b_3 & b_4 \end{bmatrix}, \quad \begin{bmatrix} 0, & 3 & 2, & 1 \\ 5, & 4, & 3, & 2 \\ 1, & 0 & 2, & 2 \end{bmatrix}, \quad \begin{bmatrix} a & b & c & d \\ a' & b' & c' & d' \\ 1 & 1 & 1 & 1 \end{bmatrix}.$$

2. What is the rank of the following matrices?

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}, \quad \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \\ 7 & 8 & 9 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Ans. 2, 2, 3, 1.

3. Solve, for $x : y : z : t$, the homogeneous equations:

$$\begin{aligned} x + y + z + t &= 0, \\ ax + by + cz + dt &= 0, \\ a^2x + b^2y + c^2z + d^2t &= 0. \end{aligned}$$

What is the rank of their matrix?

Ans. 3 if at least three of a, b, c, d are different;
1 if they are all equal; otherwise 2.

4. The complex number $2 + 3i$ consists of two linearly independent parts 2, $3i$ in the field of real numbers, but not in the complex field.

The same is true of every non-zero complex number.

5. The rational numbers $\dfrac{p}{q}$ together with all numbers such as $\dfrac{p + r\sqrt{5}}{q}$ form a field (where p, q, r are integers, and $q \neq 0$).

If five points are the vertices of a regular pentagon, prove that the ratios of all segments of all lines joining these in pairs belong to the field.

CHAPTER II

FUNDAMENTAL PROPERTIES OF THE DETERMINANT

1. Derangements.

In order to consider determinants more generally, and to make the exposition clear, we must now recall several fundamental facts of algebra. First the number of arrangements or *permutations*[1] of n different things placed in a row or column r at a time is

$$ {}_nP_r = n(n-1)(n-2)\dots(n-r+1) = \frac{n!}{(n-r)!}, \qquad (1) $$

where $n! = 1 \times 2 \times 3 \times \dots \times n$, which gives the number when all are arranged each time. Secondly, the number of *combinations*, or groups, of r things chosen from n different things is

$$ {}_nC_r = \binom{n}{r} = \frac{n!}{r!(n-r)!} = \binom{n}{n-r}. \qquad . \quad (2) $$

Thus a group of n things can be divided into subgroups, containing respectively r and $n-r$ things, in ${}_nC_r$ ways, for this is only another way of describing the same process.

Now consider the function

$$ abc \dots m $$

formed by the product of n different numbers $a, b, c, \dots, m$.

---

[1] Jacob Bernoulli (1654–1705) first used this word in this sense: *Ars conjectandi* (1713). Factorial n was introduced by Kramp (1808). The brilliant achievements of a young French mathematician, Pascal (1623–62), set this theory a-going. He discovered the number ${}_mC_n$ as $(n+1)(n+2)\dots m/(m-n)!$, and initiated the mathematical theory of probability. Moreover, it is interesting to notice that Pascal's results followed from the study of a *matrix*

$$ \begin{matrix} 1 & 1 & 1 & 1 & \dots \\ 1 & 2 & 3 & 4 & \dots \\ 1 & 3 & 6 & 10 & \dots \\ 1 & 4 & 10 & 20 & \dots \end{matrix} $$

Cf. Rouse Ball, *History of Mathematics* (London, 1901), p. 294.

Here is an example where the arrangement of factors is immaterial: all the $n!$ ways of arranging the row $a, b, \ldots, m$ are equivalent. But consider next

$$\phi = \Sigma\, a_1 b_2 c_3 \ldots m_n, \qquad . \quad . \quad . \quad . \quad (3)$$

defined as a function of n^2 numbers $a_1, a_2, \ldots, a_n, b_1, \ldots, m_n$ in which the summation sign indicates $n!$ terms, obtained by permuting $a, b, c, \ldots, m$ in all ways *without disarranging the suffixes.* This function is called the *permanent* of the square matrix of order n

$$\begin{bmatrix} a_1 & b_1 & c_1 & \ldots & m_1 \\ a_2 & b_2 & c_2 & & m_2 \\ \\ & b_n & c_n & & m_n \end{bmatrix} \qquad (4)$$

Such a function is easily constructed in any particular case. It is useful to us in paving the way for a better grasp of what a determinant is.

If $p, q, r, \ldots, t$ are the n given numbers, $a, b, c \ldots, m$, in another order, so that

$$p_1 q_2 r_3 \ldots t_n$$

is a term of the series (3), we call $pqr \ldots t$ an *inversion* or *derangement*[1] of $abc \ldots m$. We also have a special symbol

$$\begin{pmatrix} abc \ldots m \\ \cdot \ldots t \end{pmatrix}$$

to mean the *substitution* of $p, q, r, \ldots, t$ respectively for $a, b, c, \ldots, m$. It should now be clear that

$$\Sigma\, p_1 q_2 r_3 \ldots t_n$$

means exactly the same permanent as ϕ, Σ still having the same meaning.

Next, by *transposition* is meant the interchange of two of the n letters without deranging the other $n - 2$ letters. If these two letters are adjacent in the row the process is called an *adjacent transposition.* Thus

abcde, abdce, adbce, dabce, dacbe

---

[1] Cramer's word (1750).

represent terms derived by adjacent transposition in succession, whether read from left to right or right to left. Manifestly any two terms of ϕ can be connected by such a chain involving adjacent transposition, and the process can be carried out in many ways if many letters are involved.

THEOREM.—*Any transposition is equivalent to an odd number of adjacent transpositions.*

For if the transposition interchanges p and q between which k letters stand, it is equivalent to $2k + 1$ adjacent transpositions caused by shifting p through $k + 1$ places until it is just past q and then shifting q through k places back to where p first stood.

2. The C_+ and C_- Classes.

THEOREM.—*All the* n! *arrangements of* n *letters* abc . . . m *may be sorted into two classes* C_+ *and* C_-, *such that an even number of transpositions applied to any arrangement does not alter its class whereas an odd number does so. The* C_+ *class is taken to include the original arrangement* abc . . . m.

In the example just cited the terms are as follows:

C_+ *abcde, adbce, dacbe.*
C_- *abdce, dabce.*

Proof.—

This consists of two parts, first to show the *practicability* and next the *unambiguity* of the classification. First, if

$$\begin{pmatrix} 1\,2\,3\ldots n \\ i\,j\,k\ldots l \end{pmatrix}$$

denotes a substitution whereby a new arrangement $ijk \ldots l$ is derived from the n integers 1, 2, 3, . . . n, we may prove the possibility for these integers and then apply the same classification to n letters (or anything else capable of orderly arrangement). *We count how many in the lower row precede* 1, *how many greater than* 2 *precede* 2, *how many greater than* 3 *precede* 3, *and so on. Then* ijk . . . l *is placed in* C_+ *or* C_- *according as the total count is even or odd.* As this counting tallies with adjacent transpositions the classification is practicable.

Next it is unambiguous. This is proved by showing that no two arrangements $T_1 = 123 \ldots n$ and $T_r = ijk \ldots l$ are ever connected both by an odd and by an even chain of adjacent transpositions. Let the supposed chains be given by terms

$$T_1, T_2, \ldots T_{r-1}, T_r;$$
$$T_1, T_2', \ldots T_{s-1}', T_s', T_r,$$

where consecutive terms differ only by adjacent transposition. Form the chain from T_1 to T_1 by linking these chains at T_r, thus:

$$T_1, T_2, \ldots T_{r-1}, T_r, T_s', T_{s-1}', \ldots T_2', T_1.$$

If T_2 differ from its predecessor by interchange of p, q, some later term T_k differs from what immediately precedes it by interchange of q, p: otherwise the original order as in T_1 could not be finally reached. If several terms T_k have this property, we choose that nearest to T_2. Pair off T_2, T_k and start again on the first unpaired term after T_2, repeating the same argument. In this way all the series except the first T_1 is paired off. Hence either both chains are odd or both are even. This proves the theorem.

Example.
$$\begin{aligned} T_1 &= 2314 \\ T_2 &= 3214 \\ T_3 &= 3124 \\ T_4 &= 3142 \\ T_5 &= 3412 \\ T_6 &= 3421 \\ T_r = T_7 &= 4321 \\ T_8 &= 4231 \\ T_9 &= 2431 \\ T_{10} &= 2341 \\ T_{11} &= 2314 \end{aligned}$$

Here two even chains connect T_1 and T_r. We pair off, in order, rows 2, 8; 3, 6; 4, 9; 5, 11; 7, 10.

Reciprocity or Duality.

This theorem applies at once to a permanent; nor does it matter whether it is developed by fixing the suffixes and deranging the letters, or fixing the letters and deranging the suffixes. Thus when $n = 2$

$$a_1 b_2 + a_2 b_1 = a_1 b_2 + b_1 a_2.$$

Everything, in fact, that has been said of the two classes C_+ and C_- will hold of *either* suffix *or* letter permutation. But if we combine both, we obtain rather a different state of things; there would be in all $n! \times n!$ terms giving the function ϕ exactly $n!$ times, each term of ϕ occurring $n!$ times.

We may find to which class such a term as

$$p_i \quad q_j \quad r_k \ldots t_l$$

belongs by adding the total count among $pqr \ldots t$ to that of the suffixes $ijk \ldots l$, for this tallies with recovering the order $abc \ldots m$ first and then that of the suffixes $123 \ldots n$.

3. Definition of Determinant.

The series of n! *different terms*

$$\Sigma \pm a_1 b_2 c_3 \ldots m_n$$

wherein suffixes alone are permuted and the sign of the term is given by the class C_+ *or* C_- *to which it belongs is the general determinant of order* n. *It is often written*

$$(abc \ldots m),$$

or

$$|\, a_1 \quad b_2 \quad c_3 \quad \ldots \quad m_n \,|,$$

or more expressly

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 & \ldots & m_1 \\ a_2 & b_2 & c_2 & \ldots & m_2 \\ & & & & \\ a_n & b_n & c_n & & m_n \end{vmatrix}$$

The *leading diagonal* term

$$a_1 b_2 c_3 \ldots m_n$$

has a positive sign (apart of course from special negative values among its n factors).

Let $\overset{a}{\Sigma}$ denote the generation by interchanging letters, and $\underset{i}{\Sigma}$ that by interchanging suffixes. Then

$$\begin{aligned} \Delta &= \overset{a}{\Sigma} \pm a_1 b_2 c_3 \ldots m_n \\ &= \underset{i}{\Sigma} \pm a_1 b_2 c_3 \ldots m_n. \end{aligned}$$

Let us write the general term of $\sum\limits_i$

$$\epsilon\, a_i b_j c_k \ldots m_l,$$

where ϵ is ± 1 according to the class of $ijk \ldots l$. This is sometimes denoted by

$$\epsilon = \operatorname{sgn}\begin{pmatrix} 1\,2\,3 \ldots n \\ i\,j\,k \ldots l \end{pmatrix}.$$

Now the interchange of ij shows at once that

$$\operatorname{sgn}\begin{pmatrix} 1\,2\,3 \ldots n \\ j\,i\,k \ldots l \end{pmatrix}$$

is opposite to ϵ. The same remark holds of the interchange of *any* pair $i'j'$ in the lower row. Applying this to Δ, if any two letters a, c are interchanged in all terms of $\sum\limits_i$ we obtain a series Σu_N where each term u_N is equal and opposite to a term in $\sum\limits_i$. Also any two such terms u_N, $u_{N'}$ of Σu_N must differ, otherwise they would be also equal before the interchange. Thus Σu_N exactly tallies with $-\Delta$: whence

If two columns of Δ are interchanged, Δ changes sign.

For like reasons

If two rows of Δ are interchanged, Δ changes sign.

Let p, q be two letters or two suffixes. In the substitution notation these last results are written

$$\begin{pmatrix} p\,q \\ q\,p \end{pmatrix} \Delta = -\Delta.$$

Should the elements of the rows, or columns, in question, be identical, the left-hand side of this last equation leaves Δ unchanged, so that

$$\Delta = -\Delta, \qquad \text{so that } \Delta = 0,$$

whence, Δ *vanishes if two columns (or rows) are identical.*

In particular, if each element e_i of Δ is unity, the above holds, so Δ vanishes. But now, on referring to the definition, we find each term of Δ is ± 1. Hence the number of C_+ and of C_- terms is the same. But the total number is $n!$. Hence

Each class C_+ *and* C_- *has* $\frac{\mathrm{n}!}{2}$ *terms.*

An example of all possible interchanges of columns of Δ was given in (10), p. 7. There will be $n!$ ways of writing Δ in general by such interchanges of columns, together with $n!$ ways by interchanging rows. So for $n = 3$, if $\Delta = | a_1 b_2 c_3 |$

$$| a_1 b_2 c_3 | = | a_2 b_3 c_1 | = | a_3 b_1 c_2 | = - | a_1 b_3 c_2 | = - | a_2 b_1 c_3 |$$
$$= - | a_3 b_2 c_1 | = | b_2 a_1 c_3 | = \&c.$$

In this way any one of the elements may be brought to occupy the first place originally filled by a_1.

4. Arrangement of Terms in the Expansion of a Determinant. Standard Order. Co-factors.

Let us write $\Delta = T_1 + \ldots \pm T_r + \ldots \pm T_{n!}$,

where the *chief term* T_1 is $a_1 b_2 \ldots m_n$, and the number of terms is $n!$. This series is very unlike most of the familiar series of elementary analysis, for here there is real difficulty in deciding on a natural order of its terms. There is no such thing as a general nth term in Δ, as *every* term is in a sense a general term. But let us agree upon one standard order as follows. We fix the order $123 \ldots n$ of suffixes in each term, and then arrange the terms alphabetically.[1] This not only gives a unique order of succession, but automatically cuts the series of $n!$ terms into n equal sections of $(n - 1)!$ terms, exactly like the A, B, $C \ldots$ sections of a dictionary. So we write

$$\Delta = a_1 A_1 + b_1 B_1 + c_1 C_1 + \ldots + m_1 M_1, \quad . \quad (5)$$

for a_1 is a factor of all the first $(n - 1)!$ terms, b_1 of the next, and so on. The capital letter factors are called co-factors of the respective small letters. But we may equally well take the suffix order fixed as $ijk \ldots l$; so we likewise obtain, if $i = 1, 2, \ldots, n$,

$$\Delta = a_i A_i + b_i B_i + \ldots + m_i M_i, \quad . \quad . \quad (6)$$

which defines *co-factors* of elements of the ith row. Thus A_i is co-factor of a_i, B_i of b_i, and so on.

---

[1] If $n > 26$, define this as $abc \ldots za'b'c' \ldots z'a''b''c'' \ldots$ &c.!

Correlatively we may fix the letter order $abc \dots m$ of a term and write the suffix sets in ascending order, reading the suffix set as an ordinary number in a sufficiently high scale.[1] For $n = 3$ this gives

$$a_1 b_2 c_3 - a_1 b_3 c_2 - a_2 b_1 c_3 + a_2 b_3 c_1 + a_3 b_1 c_2 - a_3 b_2 c_1.$$

It leads in particular to the development

$$\Delta = a_1 A_1 + a_2 A_2 + a_3 A_3 + \dots + a_n A_n, \quad . \quad (7)$$

and in general to

$$\Delta = e_1 E_1 + e_2 E_2 + e_3 E_3 + \dots + e_n E_n \quad . \quad . \quad (8)$$

where e denotes any of the n letters.

In the above we have expanded Δ by a row or by a column. In each case *the co-factor, typified by* E_i, *is a determinant of order* $n - 1$.

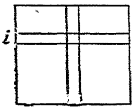
For it contains all the $(n - 1)!$ terms obtained by permuting either letters or suffixes not represented by e, or i, and the characteristic alternation of sign accompanies the derangements. Hence the definition of a determinant is satisfied.

In particular

$$A_1 = \begin{array}{cccc} b_2 & c_2 & \dots & m_2 \\ b_3 & c_3 & \dots & m_3 \\ \cdot & \cdot & \cdot\ \cdot\ \cdot & \cdot \\ b_n & c_n & \dots & m_n \end{array} \qquad (9)$$

for $a_1 A_1$ contains the term $a_1 b_2 c_3 \dots m_n$ which has the sign of the first term (the chief or leading diagonal term) in this last determinant. This is most easily remembered as the result of suppressing the row and column of Δ intersecting at a_1.

Unfortunately this last device would cause confusion in finding co-factors of other elements, because the sign of the resulting expression may be wrong. It is therefore useful to have a special name for the determinant so formed by suppressing row i and column e. It is called the *minor* of e_i. Let us represent this graphically.



---

[1] If $n < 10$ the scale of ten will do.

This process has not disturbed the orders of the rows and columns retained. Each such determinant of order $n - 1$ obtained by suppressing a row and a column of Δ is called a *first minor*. So Δ will have n^2 first minors. They are given by the elements of the matrix

$$\begin{bmatrix} +A_1, & -B_1, & +C_1, & -D_1 & \dots & (-)^{n-1}M_1 \\ -A_2, & +B_2, & -C_2, & +D_2 & \dots & \\ & & & & & \\ (-)^{n-1}A_n, & & & & \dots & +M_n \end{bmatrix} \qquad (10)$$

with alternate signs like the pattern of white and black squares on a chess board; white + and black —. For it is obvious from the mode of definition adopted for co-factor and for minor, that they only may differ in sign, whereas this sign is determined by counting the adjacent moves, say right and downwards from position A_1 to that of the element in question. Thus the minor of c_2 has a negative sign because it is three such moves from position a_1.

If n Greek letters $\alpha, \beta, \gamma, \dots, \mu$ with n suffixes are used to denote such minors we may write the determinant Δ in still more ways

$$\begin{aligned} \Delta &= a_1\alpha_1 - b_1\beta_1 + c_1\gamma_1 - \dots \\ \Delta &= a_1\alpha_1 - a_2\alpha_2 + a_3\alpha_3 - \dots \\ &\text{\&c.,} \end{aligned} \quad , \quad . \quad . \quad . \qquad (11)$$

exactly corresponding to the expansions by co-factors

$$\begin{aligned} \Delta &= a_1A_1 + b_1B_1 + c_1C_1 + \dots \\ \Delta &= a_1A_1 + a_2A_2 + a_3A_3 + \dots . \end{aligned}$$

That these are actually equivalent is best seen by noticing that the minor of b_1 is by definition $| a_2c_3d_4 \dots m_n |$, while the terms of Δ involving b_1 must be given by

$$-b_1 \, | a_2c_3d_4 \dots m_n |$$

with the negative sign, since the suffixes are in natural order while b and a alone are interchanged.

Thus *co-factors and first minors are numerically equal but differ in sign according to the scheme* (10).

An example of the use of minors with the characteristic alternate signs is given in (13), p. 7.

EXAMPLES

1. In
$$\begin{matrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{matrix}$$
the co-factor of a_1 is $b_2c_3 - b_3c_2$ which is the same as its minor. But the co-factor of b_3 is $-(a_1c_2 - a_2c_1)$, since the determinant can be written
$$-\begin{vmatrix} b_3 & a_3 & c_3 \\ b_1 & a_1 & c_1 \\ b_2 & a_2 & c_2 \end{vmatrix}$$

2. Expand (abc) by its second column; and also by its third row.

3. Prove $b_1A_1 + b_2A_2 + b_3A_3 = 0$.

4. Prove $(pa_1 + qa_2)A_3 + (pb_1 + qb_2)B_3 + (pc_1 + qc_2)C_3 = 0$.

5. What is the sign in the scheme (10) at the ith column and jth row? Ans. $(-)^{i+j}$.

5. Laplace's Development of a Determinant.

Just as the full expansion $\Sigma \pm a_1b_2 \ldots m_n$ of a determinant Δ arises from the $n!$ permutations of n different suffixes $1, 2, \ldots, n$, so also special forms of the expansion are found by considering the modified set of permutations of n things when r are alike of one kind, s of another, t of another, and so on. This number of permutations is known to be
$$\frac{n!}{r!\ s!\ t!\ldots} = \binom{n}{r,\ s,\ t,\ \ldots},$$
where $r + s + t + \ldots = n$. In particular, for two kinds, r, $n - r$, it is
$$\binom{n}{r,\ s} = \frac{n!}{r!(n-r)!} = \binom{n}{r}.$$

To fix our ideas, consider the kinds as white and black: the first r things being white, the following $n - r$ things black. Thus, as regards colour, there are only two different things, and only $\binom{n}{r}$ colour arrangements of the n original things; but using a stricter criterion, each colour arrangement subdivides into $r!(n-r)!$ different arrangements, when each individual thing is regarded as different.

Further, we may imagine all the colour arrangements of the n things first made, and then subdivided, so that we can think of the original $n!$ arrangements in their colour order; namely the first $r!(n-r)!$ of these arrangements belong to one colour order, the next $r!\,(n-r)!$ to another, and so on.

For instance, taking three things, one white w_1 and two black b_2, b_3, there are in all six arrangements

$$w_1b_2b_3,\quad w_1b_3b_2,\quad b_2w_1b_3,\quad b_3w_1b_2,\quad b_2b_3w_1,\quad b_3b_2w_1$$

derived from three colour arrangements

$$wbb,\qquad bwb,\qquad bbw.$$

For five things w_1, w_2, b_3, b_4, b_5, two white and three black, there are ten colour arrangements, say

$$wwbbb,\ \ldots,\qquad bbwbw,\qquad bbbww,$$

and twelve (2! 3!) subdivisions of each. The twelve subdivisions of *wwbbb* are

$$\left.\begin{array}{ll} w_1w_2b_3b_4b_5 & w_2w_1b_3b_4b_5 \\ w_1w_2b_3b_5b_4 & w_2w_1b_3b_5b_4 \\ w_1w_2b_4b_3b_5 & w_2w_1b_4b_3b_5 \\ w_1w_2b_4b_5b_3 & w_2w_1b_4b_5b_3 \\ w_1w_2b_5b_3b_4 & w_2w_1b_5b_3b_4 \\ w_1w_2b_5b_4b_3 & w_2w_1b_5b_4b_3. \end{array}\right\} \qquad (12)$$

Manifestly the subdivision of a given colour arrangement can be made partially, black first and then white, as in the above scheme read by columns, or white first and then black, as in the above scheme read by rows. These partial subdivisions go on independently because they each only affect arrangements entirely within a colour group.

Now let this arrangement be made of the original terms of the determinant, where the first r suffixes are called white and the next $n-r$ black, the letters $a, b, \ldots, m$ being fixed in order. We first have $\binom{n}{r}$ colour arrangements, which are next subdivided into an array of $(n-r)!$ rows and $\binom{n}{r}r!$ columns, each column containing permutations of the black but not the white.

Since the letters $a, b, \ldots, m$ are fixed in order, and the black suffixes alone are deranged in a column, each inversion being accompanied by a change of sign in the term, it follows that the sum of terms in a column is a determinant E_{n-r}, say, of order $(n-r)$ multiplied by the product of r elements whose suffixes are white. Also this determinant will appear in each of the $r!$ columns of the same original colour arrangement. Summing such $r!$ columns we obtain the sum of these coefficients, which for the same reason give a determinant E_r of order r due to permutation of white suffixes. Thus each colour arrangement gives an array of terms whose sum $E_r.E_{n-r}$ is a product of two determinants

$$\begin{aligned} E_r &= | a'_1 b'_2 \dots f'_r | \\ E_{n-} & \quad | g'_{r+1} h'_{r+2} \cdot \quad m'. \end{aligned} \tag{13}$$

where $a', b', \dots, f', g', \dots, m'$ are the letters $a, b, \dots, m$ in some order. For the letter order has been deranged by factorizing the terms of one colour grouping. Hence the original determinant is expressed as a series of $\binom{n}{r}$ terms

$$\Sigma\, E_r\, E_{n-r}$$

by rearranging the whole series of $n!$ terms in the manner explained. But inasmuch as each term of $\Sigma\, E_r.E_{n-r}$ now has its suffixes in the original order, the terms are derived from one another by applying what has been called the colour permutation to the letters $ab \dots m$ instead of to their suffixes; for this is the effect on the letters of arranging a typical term of each colour grouping (12) in ascending order of its suffixes. And finally if we examine the chief term of E_r and of E_{n-r}, which is

$$a'_1 b'_2 \dots f'_r \times g'_{r+1} h'_{r+2} \cdot \quad m'.$$

we see that it is a term of the original determinant with letters, not suffixes, deranged. We infer that, if $n > 2$,

$$a' b' \dots f' g' h' \dots m'$$

belongs with ab . . . fgh . . . m *to the* C_+ *class.* This completely specifies the sign of the term and we finally write

$$| a_1 b_2 \dots m_n | = \Sigma\, | a'_1 b'_2 \dots f'_r |\, | g'_{r+1} h'_{r+2} \dots m'. \ |, \tag{14}$$

or

$$\Delta = \Sigma\, E_r E_{n-r}.$$

This is called a *Laplace* $\left(\dfrac{\text{r}}{\text{n}-\text{r}}\right)$ *development of the determinant by its first* r *and next* n — r *rows.*[1]

Example.—

$$| a_1 b_2 c_3 d_4 | = | a_1 b_2 | \, | c_3 d_4 | + | a_1 c_2 | \, | d_3 b_4 | + | a_1 d_2 | \, | b_3 c_4 |$$
$$+ | b_1 c_2 | \, | a_3 d_4 | + | b_1 d_2 | \, | c_3 a_4 | + | c_1 d_2 | \, | a_3 b_4 | .$$

Various corollaries immediately follow. For let $ij \ldots q$ be a C_+ arrangement of the suffixes $1, 2, \ldots, n$. Then

$$\begin{aligned} | a_1 b_2 \ldots m_n | &= | a_i b_j \ldots m_q | \\ &= \Sigma \, | a'_i b'_j . \qquad . m'. \\ &= \Sigma \, E'_r \, E'_{n-r} \end{aligned}$$

say. This gives a *Laplace development by any assigned* r *rows and the complementary* n — r *rows.*

Similarly we may fix the letter order in the C_+ class and permute the suffix order. This gives a *Laplace* (r, n — r) *development by columns,* also denoted by $(r \,\vdots\, n-r)$.

Once more, by using three or more colours in the original permutations we may make *a Laplace* (r, s, t . . .) *development by rows or by columns,*

$$\Delta = \Sigma \, E_r \, E_s \, E_t \ldots, \qquad . \quad . \quad . \quad . \quad (15)$$

a series of $n!/r!\,s!\,t!\ldots$ terms.

EXAMPLES

1. Expand $| a_1 b_2 c_3 d_4 |$ as $\overset{a}{\Sigma} \, | a_i b_j | \, | c_k d_l |$ where ij, $kl = 13, 42$: also where ij, $kl = 41, 32$.

2. Expand $| a_1 b_2 c_3 d_4 e_5 |$ as $\underset{i}{\Sigma} a_1 \, | b_2 c_3 | \, | d_4 e_5 |$; also as $\Sigma a_1 \, | b_2 c_3 | \, | d_4 e_5 |$.

3. If $| a_1 b_2 c_3 d_4 e_5 f_6 |$ is expanded in various developments but without regard to the sign of the term, what sign should be attached to each of the following?

$$| b_2 d_4 | \, | a_1 c_3 e_6 f_5 |, \quad | a_6 f_5 d_4 | \, | b_3 c_2 e_1 |, \quad | a_4 b_5 c_6 | \, | d_1 e_3 f_2 | .$$

Ans. —, —, +.

4. Show that the ordinary expansion $\Sigma \pm a_1 b_2 c_3 \ldots$ of $| a_1 b_2 c_3 \ldots |$ is a particular case of a Laplace development.

5. Show that the expansion by a row or a column, e.g. $\Sigma \, a_i A_i$ is a Laplace development.

---

[1] Dating from 1772. Cf. Laplace, *Œuvres*, VIII, 365–406; Muir, *History*, I, p. 24.

6. Algebraic Complements and Minors of Order r.

If $\Delta = \Sigma\, E_r\, E_{n-r}$ is a Laplace development where each term is positive, the factors E_r, E_{n-r} are sometimes called algebraic complements of each other.

For instance in $| a_1 b_2 |$ the algebraic complement of a_2 is $-b_1$. In $| a_1 b_2 c_3 d_4 |$, that of $| a_1 c_2 |$ is $| d_3 b_4 |$.

Definition of Minor of Order r.—*The determinant obtained by suppressing any* n — r *rows and any* n — r *columns of* Δ *is called a minor of order* r.

It is also called an $(n-r)$th minor, in agreement with the first minors already introduced, where $n - r = 1$.

It is best to extend this definition so as to include, as such a minor, a determinant made by any derangement of rows and of columns. Hence we regard the two determinants given by

$$\pm | p_i\, q_j \dots s_k |$$

as minors of order r, where $p, q, \dots, s$ are any r of the n letters $a, b, \dots, m$ and $i, j, \dots, k$ are any r of the suffixes $1, 2, \dots, n$.

In this way both E_r and E_{n-r} are minors, and when their product is a term in a Laplace development of Δ, they are called *complementary minors*. Their other name, *algebraic complements*, is used also in a different sense: namely, if

$$\Delta = \Sigma\, E_r E_{n-r} = \Sigma\, | a'_1\, b'_2 \dots f'_r |\, | g'_{r+1}\, h'_{r+2} \dots m'_n |,$$

the partial letter rows $a'b' \dots f'$ and $g'h' \dots m'$ *in this order* are called algebraic complements of each other.

The phrase is used to specify two complementary letter (or suffix) sets in such cases as the Laplace development. Clearly it is relative to a given natural order. Thus, relative to the order 1234, possible algebraic complements are

1, 234	12, 34	34, 12	214, 3, &c.,
2, 314	13, 42	42, 13	
3, 124	14, 23	23, 14	
4, 213			

but not 234, 1.

7. Determinantal Permutation.

Definition.—*The* $\binom{\mathrm{n}}{\mathrm{r}}$ *arrangements of* n *letters* ab ... f, gh ... m *where* r *precede and* n — r *succeed the comma according to the* $(\mathrm{r} \vdots \mathrm{n}-\mathrm{r})$ *Laplace development is called a determinantal permutation of the* n *letters. It is denoted by*

$$\dot{a}\dot{b} \ldots \dot{f},\ \dot{g}\dot{h} \ldots \dot{m}.$$

For example we should write

$$\dot{a},\ \dot{b}\dot{c} = a,\ bc, \qquad b,\ ca, \qquad c,\ ab.$$

When this notation is used on arguments a, b, c of a function of a, b, c it is understood to mean the *sum* of the functions obtained by making these permutations.

The dot placed above a symbol (letter or suffix) indicates that it undergoes permutation. The notation evidently gives a compact way of writing a Laplace development. Thus

$$|\, a_1 b_2 c_3 d_4 e_5 \,| = |\, \dot{a}_1 \dot{b}_2 \,|\, |\, \dot{c}_3 \dot{d}_4 \dot{e}_5 \,|$$

where a series of 5!/2!3! terms is indicated. Similarly for further subdivisions, including the original expansion. So

$$\begin{aligned} \Delta = |\, a_1 b_2 c_3 d_4 e_5 \,| &= \Sigma \pm a_1 b_2 c_3 d_4 e_5 \\ &= \dot{a}_1 \dot{b}_2 \dot{c}_3 \dot{d}_4 \dot{e}_5 \\ &= |\, \dot{a}_1 \dot{b}_2 \dot{c}_3 \,|\, |\, \dot{d}_4 \dot{e}_5 \,| \\ &= |\, \dot{a}_1 \dot{b}_2 \,|\, |\, \dot{c}_3 \dot{d}_4 \,|\, \dot{e}_5, \\ &\&c. \end{aligned}$$

We may therefore speak of an $(r, s, t, \ldots)$ determinantal permutation of $r + s + t \ldots$ things. But if $r = s = t = \ldots = 1$ we must utilize a negative sign to specify certain of the terms. Thus

$$\dot{a},\ \dot{b},\ \dot{c} = a, b, c, \quad -a, c, b, \quad -b, a, c, \quad b, c, a, \quad c, a, b, \quad -c, b, a.$$

Also, as an alternative to the $\dot{a},\ \dot{b}\dot{c}$ above, we have

$$\dot{a},\ \dot{b}\dot{c} = a, bc, \quad -b, ac, \quad c, ab,$$

and so on.

As further examples of the notation we may have

$$|\dot{a}_1 b_2|\,|\dot{c}_1 d_2| = |a_1 b_2|\,|c_1 d_2| - |c_1 b_2|\,|a_1 d_2|,$$

$$\sin(A - B) = \sin \dot{A} \cos \dot{B}.$$

EXAMPLES

1. Prove $(d - c)(d - b)(d - a)(c - b)(c - a)(b - a)$

$$\begin{array}{cccc} 1 & 1 & 1 & 1 \\ a & b & c & d \\ a^2 & b^2 & c^2 & d^2 \\ a^3 & b^3 & c^3 & d^3 \end{array}$$

This determinant is called an alternant.

2. Give the corresponding identity for an alternant of orders 2, 3, and n.

3. Prove the rule of signs for the determinant $\overset{a}{\Sigma} \pm a_0 b_1 c_2 d_3$ by considering the alternant.

4. Resolve into factors

$$\left|\begin{array}{cccc} 1 & \cdot & 1 & 1 \\ a & 1 & c & d \\ a^2 & 2a & c^2 & d^2 \\ a^3 & 3a^2 & c^3 & d^3 \end{array}\right.$$

5. Prove by resolving into partial fractions that if

$$f(x) = p_0 x^r + p_1 x^{r-1} + \ldots + p_r, \qquad r < n,$$

$$\frac{f(x)}{(x - \lambda_1)(x - \lambda_2)\ldots(x - \lambda_n)}$$

$$\begin{array}{cccc} 1, & 1, & & 1 \\ \lambda_1, & \lambda_2, & & \lambda_n \\ & & & \\ \lambda_1^{n-2}, & \lambda_2^{n-2}, & \ldots, & \lambda_n^{n-2} \\ \dfrac{f(\lambda_1)}{x - \lambda_1}, & \dfrac{f(\lambda_2)}{x - \lambda_2}, & \ldots, & \dfrac{f(\lambda_n)}{x - \lambda_n} \end{array} \qquad \begin{array}{cccc} 1, & 1, & & 1 \\ \lambda_1, & \lambda_2, & & \lambda_n \\ & & & \\ \lambda_1^{n-2}, & \lambda_2^{n-2}, & \ldots, & \lambda_n^{n-2} \\ \lambda_1^{n-1}, & \lambda_2^{n-1}, & \ldots, & \lambda_n^{n-1} \end{array}$$

6. If the denominator $(x - \lambda_1)\ldots(x - \lambda_n)$ of the preceding example is written $g(x)$, express the integral of a rational function

$$\int \frac{f(x)}{g(x)}\,dx$$

as the quotient of two determinants of order n.

[Replace each $f(\lambda_i)/(x - \lambda_i)$ of row$_n$ above by $f(\lambda_i)\log(x - \lambda_i)$.

7. Evaluate $f(x)/g(x)$ and $\int\{f(x)/g(x)\}dx$ as quotients of determinants in the case when $g(x)$ has repeated factors.

[Replace col$_2$ of *both* determinants by columns obtained from col$_1$ by differentiating with regard to λ_1 as in Ex. 4. This covers the case when $\lambda_2 = \lambda_1$ alone. Further such differentiation solves the problem for higher repetition.

8. If in the alternant $\Delta = |a^0b^1c^2d^3|$ of Ex. 1, a and b are conjugate complex numbers, $r(\cos\alpha \pm i \sin\alpha)$, prove

$$: 2i \begin{vmatrix} 0 & 1 & 1 & 1 \\ r \sin\alpha & r \cos\alpha & c & d \\ r^2 \sin 2\alpha & r^2 \cos 2\alpha & c^2 & d^2 \\ r^3 \sin 3\alpha & r^3 \cos 3\alpha & c^3 & d^3 \end{vmatrix}$$

[Use $\text{col}_1 - \text{col}_2$, $\text{col}_1 + \text{col}_2$.]

9. Adapt Ex. 5 to the case when λ_1, λ_2 are conjugate complex numbers.

10. The first n integers are deranged and written also in their original order, so:

4 3 2 5 1

1 2 3 4 5

If the n pairs 11, 22, . . . nn are joined by lines, curved if necessary to avoid multiple intersections, prove that the number of intersections determines the class C_+ or C_- of the upper derangement.

[*Aitken.*

11. *Rothe's theorem on conjugate permutations.* Two derangements are conjugate if the element and place occupied in one become the place occupied and element in the other. Show that the conjugate of 43251 is 53214.

Further, show that the scheme

1 2 3 4 5

5 3 2 1 4

has the same pattern of intersection lines as in Ex. 10.

Hence prove that *two conjugate permutations belong to the same class.*

[*Aitken.*

12. A derangement is self conjugate if its conjugate is itself. Prove that its pattern is symmetrical about its horizontal bisector.

13. By considering such symmetrical patterns of n, $n - 1$ and $n - 2$ columns, prove that the relation

$$U_n = U_{n-1} + (n-1)U_{n-2}$$

connects the number of self conjugate patterns of n, $n - 1$ and $n - 2$ things respectively.

[*Rothe—Aitken.*

Cf. Muir, *History of Determinants*, **1** (1906), 60.

Count the ways in which 1 is paired with 1; then with 2; then with 3; and so on.

CHAPTER III

LINEAR PROPERTIES. FUNDAMENTAL LAPLACE IDENTITIES

1. Linearity. Homogeneity.

The n-rowed determinant $\Delta = | a_1 b_2 c_3 \ldots m_n |$ is a linear function of the elements of any row or column. So many properties of Δ hang on this that it is worth explaining in some detail.

To begin with, the notation $f(x)$ is used to denote a *function* of a *single variable* or *argument* x: while $f(x_1, x_2, x_3, \ldots x_n)$, denotes a function of n different arguments. If these arguments can usefully be called a set, or one-rowed matrix, as when they serve as co-ordinates of a point, we frequently contract this notation and write

$$f(x_i) \equiv f(x_1, x_2, x_3, \ldots, x_n). \quad . \quad . \quad . \quad (1)$$

We may even drop the suffix and write simply

$$f(x).$$

This is the *contracted functional notation* for a function of a specified set of arguments

$$x = [x_1, x_2, \ldots, x_n]. \quad . \quad . \quad . \quad . \quad . \quad (2)$$

The function is *homogeneous* and of *order* p in its arguments if, and only if,

$$k^p f(x_1, x_2, \ldots, x_n) = f(kx_1, kx_2, \ldots, kx_n) \qquad (3)$$

identically for all values of k. Thus if $a_1, a_2, \ldots, a_n$ are independent of x, the function

$$a_1 x_1 + a_2 x_2 + \ldots + a_n x_n \quad . \quad . \quad . \quad (4)$$

is of order unity. It is a *linear homogeneous form* in n arguments x. We write this in various ways, for example,

$$\sum_{i=1}^{n} a_i x_i, \qquad (a|x), \qquad a_x, \qquad (ax). \quad . \quad . \quad (5)$$

Now the fundamental property of a linear form is the simplicity of its addition theorem, namely

$$f(x+y)=f(x)+f(y),$$

which can be interpreted in the contracted functional notation. Thus if

$$f(x)=a_1x_1+a_2x_2+\ldots+a_nx_n$$
$$f(y)=a_1y_1+a_2y_2+\ldots+a_ny_n$$

then

$$f(x+y)=a_1(x_1+y_1)+a_2(x_2+y_2)+\ldots+a_n(x_n+y_n).$$

More generally if we multiply throughout by p and q respectively and add,

$$f(px+qy)=pf(x)+qf(y)$$

where

$$px+qy=px_1+qy_1,\ px_2+qy_2,\ \ldots,\ px_n+qy_n.$$

An immediate consequence is the following theorem.

The n*-rowed determinant* Δ *is unaltered in value by adding to one of its columns any linear combination of its other columns. This is true also of its rows.*

Thus, by p col$_2$ + q col$_3$ + r col$_4$,

$$\begin{vmatrix} a_1+pb_1+qc_1+rd_1, & b_1, & c_1, & d_1 \\ a_2+pb_2+qc_2+rd_2, & b_2, & c_2, & d_2 \\ a_3+pb_3+qc_3+rd_3, & b_3, & c_3, & d_3 \\ a_4+pb_4+qc_4+rd_4, & b_4, & c_4, & d_4 \end{vmatrix}$$
$$=(abcd)+p(bbcd)+q(cbcd)+r(dbcd)$$
$$=(abcd),$$

because the other terms, each having an identical pair of columns, vanish. This proves Cramer's rule of p. 11. For if

$$a_ix+b_iy+c_iz=d_i,\qquad i=1,2,3,$$

then $\quad (dbc)=(ax+by+cz\,.\,bc)$

$$=(abc)x+(bbc)y+(cbc)z=(abc)x\ \text{ only}.$$

Consequently $x=(dbc)/(abc)$ provided that $(abc)\neq 0$. Similarly y is calculated from (adc); and so on for z and for n such equations.

2. Special Determinants.

The following cases, the results of which can be easily verified, are worth noticing.

(1) *The unit determinant.* It has unity for each leading diagonal element, and zero for each other element.

$$1 = \begin{vmatrix} 1 & \\ & 1 \end{vmatrix} \quad \begin{vmatrix} 1 & \\ & 1 \end{vmatrix} = \&c.$$

The matrix of this determinant is called the unit matrix (§3, p. 68).

(2) The *value* of the unit determinant is unaltered by filling up one triangle of zeros with arbitrary elements.

$$1 = \begin{vmatrix} 1 & . \\ x & 1 \end{vmatrix} = \begin{vmatrix} 1 & . & . \\ x & 1 & . \\ y & z & 1 \end{vmatrix} = \&c.$$

(3)
$$\begin{matrix} a_1 & . & . & . \\ x & b_2 & . & . \\ y & z & c_3 & . \\ p & q & r & d_4 \end{matrix} = a_1 b_2 c_3 d_4 .$$

(4) A determinant of lower order can be expressed as one of higher order without disarranging its elements.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & b & x \\ c & d & y \\ . & . & 1 \end{vmatrix} = \begin{vmatrix} a & b & x & z \\ c & d & y & t \\ . & . & 1 & . \\ . & . & . & 1 \end{vmatrix} = \begin{vmatrix} a & b & . & . \\ c & d & . & . \\ e & f & 1 & . \\ g & h & . & 1 \end{vmatrix}$$

Thus the determinant $\begin{vmatrix} a & b \\ c & d \end{vmatrix}$ is extended diagonally with the unit matrix, bordered on one side with zeros and on the other with *arbitrary* elements.

3. Double Suffix Notation and other Contractions.

Hitherto we have used letters to distinguish columns, and suffixes for rows. This has certain advantages, but not such as entirely supersede other notations. Let us now write

$$a_1 \quad a_2 \quad a_3 \quad \ldots \quad a_n$$

for

$$a \quad b \quad c \quad \ldots \quad m$$

and a_{ij} for the letter of the ith row and the jth column: so that a typical determinant is

$$\Delta = \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{array} = |\, a_{ij} \,| \qquad (6)$$

We adopt this simple notation[1] $|\, a_{ij} \,|$ for the determinant Δ, and $[a_{ij}]$ for its matrix, so that

$$[a_{ij}] = \begin{array}{ccc} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{array} \quad . \qquad (7)$$

In consonance with a previous use, Δ is sometimes denoted by $(a_1 a_2 a_3)$, where a_i stands for the ith column.

As a rule there is no ambiguity in practice when the order n of the determinant is unspecified, so that $|\, a_{ij} \,|$ is of whatever order immediately concerns us. Where doubt may exist the order must be clearly explained.

A particular case of this notation is defined by

$$|\, \delta_{ij} \,|, \ [\delta_{ij}]; \qquad \delta_{ij} = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases} \qquad (8)$$

This symbol, which is called the Kronecker delta, characterizes the *unit determinant* and the *unit matrix*.

4. A Determinant is irresoluble into Factors.

Regarded as a rational integral function of its n^2 elements, a determinant has no rational factors. For suppose if possible that $\Delta = |\, a_{ij} \,|$ can be written as the product of two rational factors $\theta\phi$.

Since Δ is linear in each element, a_{11} cannot occur in both factors θ, ϕ. Suppose that it occurs in θ.

In the expansion of the determinant no term occurs in which

---

[1] Introduced by H. J. S. Smith (1862) and established by Kronecker.

a_{11} is multiplied by any element belonging to its row or column. Thus ϕ can involve no element belonging to the first row or the first column. Let a_{rs} be an element which does occur in ϕ. By similar reasoning no element belonging to the rth row or sth column can occur in θ.

Thus the two elements a_{r1}, a_{1s} cannot occur either in θ or in ϕ. But the expansion of the determinant involves every element. Our supposition that Δ can be written as a product of factors $\theta\phi$ is therefore untenable.

5. Rules for Combining Matrices.

It is now the place to give the rules for addition and subtraction of matrices. These rules, which are due to Cayley, turn out to justify themselves, although they contradict some of the corresponding rules for determinants.

If a_{ij} and b_{ij} are corresponding elements in row i and column j of two matrices A and B, *the sum of* A *and* B *is a matrix with* $a_{ij} + b_{ij}$ *for corresponding element.* This is the definition on the understanding that it is true for all values of i and j, so that A and B must be *conformable for addition*; they must each have the same number n of columns and m of rows.

A rule is sometimes given for addition, when the matrices are unconformable; but this case will not be considered.

Let the sign $\{ij\}$ placed after an equality mean "identically for all values of i and j"; then the m by n matrix C is the sum of A and B if

$$c_{ij} = a_{ij} + b_{ij} \qquad \{ij\} \quad . \quad . \quad . \quad . \quad (9)$$

where $A = [a_{ij}]$, $B = [b_{ij}]$, $C = [c_{ij}]$. We now write this comprehensively as

$$C = A + B. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Likewise for subtraction, we define

$$C = A - B \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

to mean

$$c_{ij} = a_{ij} - b_{ij} \qquad \{ij\}. \quad . \quad . \quad . \quad . \quad (12)$$

In particular we write $A + A = 2A$, so that $2A$ denotes a matrix wherein *each* element of A is doubled: whence if r is a positive integer

$$rA = [ra_{ij}] \qquad . \quad . \quad . \quad . \quad . \quad (13)$$

where every element is multiplied by r.

Again, if $A = B$ then $a_{ij} = b_{ij}$ $\{ij\}$, while if $A + B$ is the null matrix, $a_{ij} + b_{ij} = 0$. Accordingly we write

$$[a_{ij}] - [a_{ij}] = [a_{ij} - a_{ij}] = [0] = 0:$$

also

$$[a_{ij}] + [-a_{ij}] = [a_{ij} - a_{ij}] = 0.$$

This in fact defines $-A$, namely

$$A = [a_{ij}], \quad -A = -[a_{ij}] = [-a_{ij}]. \qquad (14)$$

The reader who has examined the theory of indices in elementary algebra will have no difficulty in extending the validity of relation (13) to cover cases where r is not merely a positive integer, but is negative (as in (14)), zero, rational, real or complex. Let us call such values *scalar* numbers to distinguish them from the entities A, B, C which are *arrays* of numbers, although they behave in many ways like scalar or ordinary numbers. This behaviour is summed up by saying:

Linear combinations of matrices with scalar coefficients obey the rules of ordinary algebra.

In fact we may prove without difficulty the following fundamental identities: $A = B$ implies $B = A$,

$$\begin{aligned} A + B &= B + A, \\ (A + B) + C &= A + (B + C), \\ rA + rB &= r(A + B), \\ rA + sA &= (r + s)A, \\ rA &= Ar. \end{aligned} \qquad . \quad (15)$$

Each of these equations involving A, B, C is merely an abbreviation for a set of mn equations involving a_{ij}, b_{ij}, c_{ij}, where both i and j remain constant in each particular equation.

We may even have relations linear in the matrices but not linear in scalar numbers. It would be true to say

$$\frac{xA}{y - z} + \frac{yB}{z - x} = \frac{x(z - x)A + y(y - z)B}{(y - z)(z - x)}$$

where x, y, z are scalar numbers. What is at present excluded is a product of matrices AB, AC, A^2, . . . , which will later be defined.

The reader who is familiar with the use of vectors in one form or another will recognize that these laws are identical with the addition laws of vectors.

Transposition of a Matrix.

Definitions.—*The matrices* $A = [a_{ij}]$ *and* $A' = [a_{ji}]$ *are called the transposed of one another. Each is obtained from the other by interchanging its rows and columns. It is often convenient to denote the transposed of* A *by an accent.* (Cf. Chap. I, §3.)

The property is conjugate or symmetrical, and sometimes A' is called the conjugate of A.

When transposition leaves a matrix unaltered, the matrix is said to be *symmetrical:* if transposition is equivalent to changing the sign of all the elements the matrix is *skew symmetrical.*

When a matrix has a single row, or a single column, it is called a *vector.* Thus there are two distinct types of vector, the row vector, and the column vector.

EXAMPLES

If $A \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 2 & 3 \\ 4 & 4 & 6 \\ 7 & 8 & 8 \end{bmatrix}$,

then

$$A + B = \begin{bmatrix} 1 & 4 & 6 \\ 8 & 9 & 12 \\ 14 & 16 & 17 \end{bmatrix}, \text{ while } A - B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Again

$$pA + qB = \begin{bmatrix} p \quad , & 2p + 2q, & 3p + 3q \\ 4p + 4q, & 5p + 4q, & 6p + 6q \\ 7p + 7q, & 8p + 8q, & 9p + 8q \end{bmatrix}.$$

If A' is the transposed of A, then $A' = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$.

2. $Q = \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$ is symmetrical, while $S = \begin{bmatrix} 0 & r & q \\ -r & 0 & p \\ -q & -p & 0 \end{bmatrix}$ is skew symmetric.

A skew symmetric matrix necessarily has zero elements throughout the leading diagonal.

3. Prove that in general the sum of a square matrix and its transposed matrix is *symmetrical,* while the difference is *skew symmetrical.*

4. Prove the determinant of the matrix equivalent to pA is p^n times the determinant of A, if A is a square matrix of order n.

5. If I denotes the *unit matrix* $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, then $A - \lambda I$ is

$$\left[\begin{array}{ll} a_{11} - \lambda, & a_{12} \quad , \\ a_{21} \quad , & a_{22} - \lambda, \end{array}\right.$$

The determinant $| A - \lambda I |$ is a polynomial of degree n in λ. Likewise for $| A + \lambda B |$ if B is nonsingular, square and of order n.

6. Prove $(A + B)' = A' + B'$.

7. Prove $A = (A')'$.

8. Any square matrix of order n has at most n^2 arbitrary elements. The symmetrical has $\frac{1}{2}n(n + 1)$ and the skew symmetrical $\frac{1}{2}n(n - 1)$ arbitrary elements.

6. Currency of a Matrix.

It is often very well worth while to group several columns as well as rows of a determinant in one symbol. Let us agree to use capital letters with suffixes for this purpose, unless something is said to the contrary.

We first write the n-rowed determinant

$$(a_1 a_2 a_3 \ldots a_r b_1 b_2 b_3 \ldots b_{n-r}) \quad . \quad . \quad (16)$$

as

$$(A_r B_{n-r}). \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

In full this is given by

$$\Delta = \left| \begin{array}{llllllll} a_{11} & a_{12} & . \; . & a_{1r} & b_{11} & b_{12} & . \; . & b_{1s} \\ a_{21} & a_{22} & & a_{2r} & b_{21} & b_{22} & & \\ & & & & & & & \\ a_{n1} & a_{n2} & . \; . & a_{nr} & b_{n1} & b_{n2} & . \; . & b_{n} \end{array} \right. \qquad (18)$$

where $s = n - r$. Clearly this is a formidable expression, only to be used sparingly, while (16) is much easier to handle and (17) is even better still. In (16) each a_i or b_i denotes a column; in (17) A_r, B_{n-r} denote complementary oblong matrices making together the square which furnishes determinant (18).

Definition of Currency.—*The suffix* r *of* $\mathrm{A_r}$ *is the currency of the matrix* $[\mathrm{A_r}]$ *for the field or category of order* n.

So the currency specifies the number of columns in A_r. In the same way we consider each a_i, b_i of (16) to have unit currency. They act as "small change" equivalent to two "pieces" A_r, B_{n-r} of higher currency in the determinant whose contents is of total currency n.

Next let Δ be cut just below the rth row and expanded by Laplace's development. The result is a sum of $\binom{n}{r}$ terms

$$\Sigma (a_1 a_2 a_3 \ldots a_r)_{123\ldots r} \cdot (b_1 b_2 b_3 \ldots b_s)_{r+1 \cdot r+2 \ldots n}$$

where $a_1, a_2, \ldots, b_s$ are deranged, although the outside suffixes are fixed because they refer to rows. It is essential to have a ready way of alluding to this fundamental operation, so we simply denote the sum of all the $\binom{n}{r}$ arrangements by these equivalent notations: either $\dot{A}_r$, $\dot{B}_s$ or $\dot{a}_1 \dot{a}_2 \ldots \dot{a}_r$, $\dot{b}_1 \dot{b}_2 \ldots \dot{b}_s$. This is a case of what has been called (p. 27) a *determinantal permutation* of the r columns of the matrix A_r with the s columns of B_s.

7. Transposition Properties of Determinants.

The process just described can be carried further by partitioning one or both of A_r and B_s. Equally well we may partition a determinant into layers of rows, or even make a double partition by columns and rows. In this case we express a determinant by the matrices in the rectangular partitions. For example

$$\begin{array}{cccc} a & b & & \\ e & f & & \\ p & q & & \\ x & y & z & t \end{array}$$

can be looked on as pieced together from four sub-matrices in rectangular array

$$\begin{array}{cc} \begin{bmatrix} a & b \\ e & f \end{bmatrix} & \begin{bmatrix} c & d \\ g & h \end{bmatrix} \\ \begin{bmatrix} p & q \\ x & y \end{bmatrix} & \begin{bmatrix} r & s \\ z & t \end{bmatrix} \end{array}$$

forming a matrix of matrices. In all such partitioning the relative positions of the original elements of the determinant are maintained. Accordingly by

$$\Delta = \left| \begin{array}{cc} A & B \\ D & E \end{array} \right. \tag{19}$$

is meant a determinant whose elements are partitioned into four matrices A, B, D, E. If A, D each have r columns, B, E, s columns, A, B, t rows, D, E, u rows, then

$$r+s=t+u=n,$$

Δ being of order n. Now we can alter the expression for Δ in various useful ways. For by transposing all s columns of B, E to precede columns of A, D we have

$$\Delta=\begin{vmatrix} A & B \\ D & E \end{vmatrix}=(-)^{rs}\begin{vmatrix} B & A \\ E & D \end{vmatrix}. \quad . \quad . \quad . \quad (20)$$

By transposing rows, we have also

$$\Delta=(-)^{tu}\begin{vmatrix} D & E \\ A & B \end{vmatrix}=(-)^{rs+tu}\begin{vmatrix} E & D \\ B & A \end{vmatrix}.$$

Similarly for triple and higher partitions.

The main use of this matrix notation occurs when *all the matrices have* n *rows*, and determinants of orders $2n$, $3n$, . . . , pn are considered.

For example, let R, S, T be matrices of n rows and r, s, t columns respectively, where

$$r+s+t=2n. \quad . \quad . \quad . \quad . \quad . \quad (21)$$

Then

$$\Delta=\begin{vmatrix} R & S & . \\ R & . & T \end{vmatrix}$$

represents a determinant of $2n$ rows and $2n$ columns, the dots signifying arrays of zeros. Now Δ is unaltered by

$$\text{row}_{n+1}-\text{row}_1, \quad \ldots \quad \text{row}_{n+i}-\text{row}_i, \quad \ldots \quad \text{row}_{2n}-\text{row}_n,$$

or briefly

lower matrix — upper matrix.

But by definition of subtraction of matrices

$$R-R=0, \quad 0-S=-S, \quad T-0=T.$$

Thus

$$\Delta=\begin{vmatrix} R & S & . \\ . & -S & T \end{vmatrix}. \quad . \quad . \quad . \quad . \quad (22)$$

Similarly

$$\Delta=\begin{vmatrix} . & S & -T \\ R & . & T \end{vmatrix}.$$

Further, on multiplying by -1 each of the last n rows of the full expression summarized in (22), we change the sign of all the elements concerned, and so $-S$ becomes $+S$ and T becomes $-T$. Hence

$$\Delta = (-)^n \begin{vmatrix} R & S & \\ & S & -T \end{vmatrix}$$

But T has t columns. So, on multiplying the last t columns of Δ by -1, we obtain

$$\begin{aligned} \Delta &= (-)^{n+t} \begin{vmatrix} R & S & . \\ . & S & T \end{vmatrix}, \\ &= (-)^{n+t+rs} \begin{vmatrix} S & R & . \\ S & . & T \end{vmatrix} \quad . \quad . \quad . \quad . \quad (23) \end{aligned}$$

as in (20).

So the determinants $\begin{vmatrix} R & S & \\ R & . & T \end{vmatrix}$ and $\begin{vmatrix} S & R & . \\ S & . & T \end{vmatrix}$ can only differ in sign, and that only when $n + t + rs$ is odd.

Exactly similar reasoning shows that if

$$R, L, M, N, \ldots$$

are p matrices of currency $h, i, j, k \ldots$ respectively, the determinant of order np

$$\begin{vmatrix} R & L & . & . \\ R & . & M & . \\ R & & & N \end{vmatrix} \qquad (24)$$

apart from sign, is unaltered by writing any matrix L or M or $N \ldots$ repeated in the first column and the rest diagonally as before. The only condition is

$$h + i + j + k + \ldots = (p-1)n \quad . \quad . \quad . \quad (25)$$

to make the total number of *elements* (not matrices) the same in each row and column.

The next section will illustrate this type of determinant.

EXAMPLES

1. If $A, B, C, \ldots$ are each two by two (or $2k$ by $2k$) matrices prove

$$\begin{vmatrix} A & B \\ C & D \end{vmatrix} = \begin{vmatrix} D & C \\ B & A \end{vmatrix} = \&c.$$

2. If $A, B, C, \ldots$ are each square matrices, then

$$\begin{vmatrix} A & . & . \\ . & B & . \\ . & . & C \end{vmatrix} = |A|\ |B|\ |C|.$$

3. Prove

$$\begin{vmatrix} A & . & . \\ L & B & . \\ M & N & C \end{vmatrix} = |A|\ |B|\ |C|.$$

[Expand by Laplace's method.]

4. If

$$\begin{array}{cccccc} a_1 & b_1 & c_1 & l_1 & m_1 & x_1 \\ a_2 & b_2 & c_2 & l_2 & m_2 & x_2 \\ a_3 & b_3 & c_3 & l_3 & m_3 & x_3 \\ \\ a_6 & b_6 & c_6 & l_6 & m_6 & x_6 \,| \end{array} \equiv (A_3 L_2 X_1) \equiv (ALX),$$

show that $(ALX) = (AXL) = -(XAL) = -(XLA) = (LAX) = -(LXA)$.

5. Examine the corresponding six permutations for $(A_i\, L_j\, X_k)$, $i + j + k = n$.

6. Extend the linearity theorem at the middle of p. 31, to the case when the letters a, b, c, d are replaced by matrices.

8. Fundamental Laplace Identities.

Laplace's development leads to many important results in particular cases, some of which will now be given. They mostly depend on cutting the determinant half-way across and expanding by complementary minors of equal order.

First expand the vanishing determinant

$$\left|\begin{array}{cc:cc} x_1 & y_1 & x_1 & y_1 \\ x_2 & y_2 & x_2 & y_2 \\ x_3 & y_3 & x_3 & y_3 \\ x_4 & y_4 & x_4 & y_4 \end{array}\right|$$

then

$$(xy)_{12}\ (xy)_{34} + (xy)_{13}\ (xy)_{42} + (xy)_{14}\ (xy)_{23} = 0 \qquad (26)$$

identically.

Correlatively since

$$\begin{array}{cccc} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \end{array} = 0,$$

then

$$(bc)\ (ad) + (ca)\ (bd) + (ab)\ (cd) = 0. \qquad (27)$$

Several results follow from a determinant of the sixth order. Consider the identity

$$\left|\begin{array}{cccccc} a_1 & b_1 & c_1 & d_1 & \cdot & \cdot \\ a_2 & b_2 & c_2 & d_2 & \cdot & \cdot \\ a_3 & b_3 & c_3 & d_3 & \cdot & \cdot \\ \hdashline \cdot & \cdot & \cdot & d_1 & e_1 & f_1 \\ \cdot & \cdot & \cdot & d_2 & e_2 & f_2 \\ \cdot & \cdot & \cdot & d_3 & e_3 & f_3 \end{array}\right| = \left|\begin{array}{cccccc} a_1 & b_1 & c_1 & d_1 & \cdot & \cdot \\ a_2 & b_2 & c_2 & d_2 & \cdot & \cdot \\ a_3 & b_3 & c_3 & d_3 & \cdot & \cdot \\ \hdashline -a_1 & -b_1 & -c_1 & \cdot & e_1 & f_1 \\ -a_2 & -b_2 & -c_2 & \cdot & e_2 & f_2 \\ -a_3 & -b_3 & -c_3 & \cdot & e_3 & f_3 \end{array}\right|$$

This is comprised in the single operation: *Subtract the upper from the lower half matrix in the first determinant.* Expanding by Laplace's method we have

$$(abc)\,(def) = (dbc)\,(aef) + (adc)\,(bef) + (abd)\,(cef), \qquad (28)$$

where all suffixes are 1, 2, 3. All the other usual terms have disappeared because of the zero columns in one or other factor.

This is called a *fundamental ternary identity*, in general form. But in particular if $[d] = [f]$ (i.e. $d_1 = f_1$, $d_2 = f_2$, $d_3 = f_3$) then $(def) = 0$ and

$$0 = (fbc)\,(aef) + (afc)\,(bef) + (abf)\,(cef).$$

Rearranged this is

$$0 = (bcf)\,(aef) + (caf)\,(bef) + (abf)\,(cef). \qquad (29)$$

This is an example of an *extensional*, for it reproduces at a higher order an identity (27) already known, merely by inserting a common letter f in each factor.

The eighth order gives still more varied results by exactly the same device. We expand the identity

$$\left|\begin{array}{cccccccc} a_1 & b_1 & c_1 & d_1 & e_1 & f_1 & \cdot & \cdot \\ a_2 & b_2 & c_2 & d_2 & e_2 & f_2 & \cdot & \cdot \\ a_3 & b_3 & c_3 & d_3 & e_3 & f_3 & \cdot & \cdot \\ a_4 & b_4 & c_4 & d_4 & e_4 & f_4 & \cdot & \cdot \\ \hdashline \cdot & \cdot & \cdot & d_1 & e_1 & f_1 & g_1 & h_1 \\ \cdot & \cdot & \cdot & d_2 & e_2 & f_2 & g_2 & h_2 \\ \cdot & \cdot & \cdot & d_3 & e_3 & f_3 & g_3 & h_3 \\ \cdot & \cdot & \cdot & d_4 & e_4 & f_4 & g_4 & h_4 \end{array}\right| = \left|\begin{array}{cccccccc} a_1 & b_1 & c_1 & d_1 & e_1 & f_1 & \cdot & \cdot \\ a_2 & b_2 & c_2 & d_2 & e_2 & f_2 & \cdot & \cdot \\ a_3 & b_3 & c_3 & d_3 & e_3 & f_3 & \cdot & \cdot \\ a_4 & b_4 & c_4 & d_4 & e_4 & f_4 & \cdot & \cdot \\ \hdashline -a_1 & -b_1 & -c_1 & \cdot & \cdot & \cdot & g_1 & h_1 \\ -a_2 & -b_2 & -c_2 & \cdot & \cdot & \cdot & g_2 & h_2 \\ -a_3 & -b_3 & -c_3 & \cdot & \cdot & \cdot & g_3 & h_3 \\ -a_4 & -b_4 & -c_4 & \cdot & \cdot & \cdot & g_4 & h_4 \end{array}\right|$$

and obtain

$$(abcd)\,(efgh) + (abce)\,(fdgh) + (abcf)\,(degh)$$
$$= (adef)\,(bcgh) + (bdef)\,(cagh) + (cdef)\,(abgh). \quad . \quad (30)$$

All other terms disappear because of zero columns. Now it is useful to use the abbreviation as in § 7, p. 27, for this last result. It is written

$$(abc\dot{d})\,(\dot{e}\dot{f}gh) = (\dot{a}def)\,(\dot{b}\dot{c}gh), \quad . \quad . \quad . \quad (31)$$

where the three dots placed over the letters of a term indicate the sum of three terms obtained by suitable derangement—in one case

$$d,\ ef \qquad e,\ fd \qquad f,\ de,$$

and in the other,

$$a,\ bc \qquad b,\ ca \qquad c,\ ab.$$

Matrix Notation.

As this eight-rowed determinant can teach us several further interesting facts about four-rowed determinants, let us make a natural abbreviation, writing

$$\begin{array}{cccccccc} a & b & c & d & e & f & . & . \\ . & . & . & d & e & f & g & h \end{array} \qquad \begin{array}{cccccccc} a & b & c & d & & f & & \\ -a & -b & -c & & & & & h \end{array}$$

as short for the above equality. In this last each letter stands for a matrix of one column of four elements. This enables us to form other possible relations by varying the number of repeated letters. Thus from

$$\begin{array}{cccccccc} a & b & c & d & e & . & . & . \\ . & . & c & d & e & f & g & h \end{array} \qquad \begin{array}{cccccccc} a & b & c & d & e & . & . & . \\ -a & -b & . & . & . & f & g & h \end{array} \quad . \quad (32)$$

it follows that

$$(abcd)\,(efgh) + (abde)\,(cfgh) + (abec)\,(dfgh)$$
$$= (acde)\,(bfgh) - (bcde)\,(afgh), \quad . \quad (33)$$

or, more succinctly,

$$(ab\dot{c}\dot{d})\,(\dot{e}fgh) = (\dot{a}cde)\,(\dot{b}fgh). \quad . \quad . \quad (34)$$

The five dots placed above letters, indicating determinantal permutation as already explained, must, of course, not be confused with the dots in (32) which stand for zeros. The two uses happen to have arisen quite independently, like the use of vertical lines $|a_{ij}|$ to denote a determinant or else a modulus of a complex number. It is an interesting study to discover how mathematics advances by the bold use of one symbol for two or more distinct things, letting the context decide.

Comparing (32) and (34) a curious rule comes to light. The dots used in the two statements nicely balance each other. Letters c, d, e in (32) are dotted in (34). This should facilitate the proof of any such identity.

Examples.—Prove

$$(\dot{a}\dot{b}c)\,(\dot{d}\dot{e}f) = 0,$$

$$(\dot{a}\dot{b}\dot{c}\dot{d})\,(\dot{e}fgh) = 0,$$

$$(\dot{a}\dot{b}cd)\,(\dot{e}\dot{f}gh) = (abef)\,(cdgh),$$

$$(\dot{a}\dot{b}cdu)\,(\dot{e}\dot{f}ghu) = (abefu)\,(cdghu).$$

9.[1] Fundamental Identities of Order n.

Let $i + j + k = n$, so that

$$(A_i B_j C_k)$$

denotes an n-rowed determinant whose columns are specified by

$$a_1, a_2, \ldots, a_i, b_1, b_2, \ldots, b_j, c_1, c_2, \ldots, c_k$$

in this order. Let

$$(D_i E_j F_k)$$

be another such determinant whose columns are specified by

$$d_1, d_2, \ldots, d_i, e_1, e_2, \ldots, e_j, f_1, f_2, \ldots, f_k.$$

We form the determinant of $2n$ rows

$$\begin{array}{cccccc} A_i & B_j & C_k & D_i & . & . \\ . & . & C_k & D_i & E_j & F_k \end{array}$$

or simply

$$\begin{array}{cccccc} A & B & C & D & . & . \\ . & . & C & D & E & F \end{array}\Bigg|$$

meaning exactly the same thing. Now we subtract the upper n rows from the lower, as before, so that

---

[1] The rest of this chapter, excepting §12, may be omitted on a first reading.

$$\left|\begin{array}{cccccc} A & B & C & D & . & . \\ . & . & C & D & E & F \end{array}\right. \qquad \begin{array}{cccccc} A & B & C & D & . & . \\ -A & -B & . & . & E & F \end{array} = (-)^i \begin{array}{cccccc} D & B & C & A & . & . \\ . & -B & . & -A & E & F \end{array}$$

on interchanging i columns D with i columns A. Expanding the first and third of these by a Laplace development, as in §8, p. 42, we obtain a sum of $\binom{2n}{n}$ terms on each side of the identity; but many of these terms vanish, because of zero columns. On the left the surviving terms, $\binom{k+i}{k}$ in number, may be written

$$(A_i B_j \dot{C}_k)\,(\dot{D}_i E_j F_k).$$

On the right the upper row of capitals furnishes the first factor and the lower the second factor of a term. Each second factor has exactly i negative columns. Hence the expansion is

$$(-)^{2i}\,(D_i \dot{B}_j C_k)\,(\dot{A}_i E_j F_k)$$

with $\binom{i+j}{i}$ terms. Equating these results we have the important identity

$$(A_i B_j \dot{C}_k)\,(\dot{D}_i E_j F_k) = (D_i \dot{B}_j C_k)\,(\dot{A}_i E_j F_k). \qquad \text{(I)}$$

In particular if $j = 0$, $i + k = n$, we merely suppress the matrices B and E, so that

$$(A_i \dot{C}_k)\,(\dot{D}_i F_k) = (D_i C_k)\,(A_i F_k). \quad . \quad . \quad \text{(II)}$$

with only one term on the right, and $\binom{i+k}{i}$ terms on the left.

Once more, from the equalities

$$\begin{array}{cccccc} A_i & B_j & C_k & D_i & E_j & . \\ . & B_j & C_k & D_i & E_j & F_k \end{array} \qquad \begin{array}{cccccc} A & B & C & D & E & . \\ . & B & C & D & E & F \end{array} \qquad \begin{array}{cccccc} A & B & C & D & E & . \\ -A & . & . & . & . & F \end{array}$$

we have on expansion

$$(A_i \dot{B}_j \dot{C}_k)\,(\dot{D}_i \dot{E}_j F_k) = 0. \quad . \quad . \quad . \quad \text{(III)}$$

These three results, which are of great use in the invariant theory, can be summed up in one statement:

A sum of products of two n*-rowed determinants, formed by determinantal permutation of* p *columns of one with* q *columns* *of the other, is identically equal to a like sum, a single product, or zero, according as* p + q $<, =, >$ n. *All columns undergoing derangement on one side of the identity are collected into one factor on the other side, except in the third alternative case.*

Letters which are so collected into one factor are said to be *convolved* in that factor.

This theorem, which collects together results illustrated in §8, has only comparatively lately been recognized, although its simpler cases[1] when $n = 2, 3, 4$, &c., were studied from the very outset of the determinant theory. Sylvester was the first to establish it as in formula (II), but he did not arrive at the other cases (I) and (III).

A particular case of (II) when $i = 1$ can be written

$$(a_1 b_2 b_3 \ldots b_n)(b_1 D) - (a_1 b_1 b_3 \ldots b_n)(b_2 D) + \ldots = (b_1 b_2 \ldots b_n)(a_1 D)$$

by substituting a_1 for A_i, $b_2 b_3 \ldots b_n$ for C_k, b_1 for D_i, and D for F_k. This falls in with the Cramer rule of substitutions if we rewrite it as

$$\begin{aligned}(a_1 b_2 b_3 \ldots b_n)(b_1 D) + (b_1 a_1 b_3 \ldots b_n)(b_2 D) + \cdot\ \cdot \\ + (b_1 b_2 \ldots b_{n-1} a_1)(b_n D) = (b_1 b_2 \ldots b_n)(a_1 D). \qquad \text{(IV)}\end{aligned}$$

This last result is closely connected with the easily proved identity (V) given later, p. 50, where examples of its use will be found.

10. Implicit and Explicit Convolution.

In the above identities certain matrices are unchanged for each term. For instance, $A_i B_j$ on the left of (I) remains unpermuted while $E_j F_k$ is unchanged on both sides of the identity. It is useful to make the following distinction and to say that $A_i B_j$ is *explicitly convolved* on the left and *implicitly convolved*[2] on the right, while $E_j F_k$ is explicitly convolved throughout. Similarly $C_k D_i$ is explicitly convolved on the right and implicitly on the left.

---

[1] In 1779 Bézout gave several simple cases. Cf. Muir, *History of Determinants*, Vol. I, p. 41.

[2] The importance of a word for this implicit convolution has begun to be felt elsewhere. Such matrices are "herausgegriffenen". Weitzenböck, *Math. Annalen*, **97** (1927), 794.

These identities are given in the *Trans. Cambridge Phil. Soc.* XXI (1909), under identity B, p. 209, where incidentally there is an error in the sign, which should read $(-)^{s(q-1)+1}$ and not $(-)^{sq}$. Identity (IV) was formulated by Sylvester in 1839, and (II) in 1851 (*Phil. Mag.* (4), ii, 142–145).

11. General Fundamental Identities of Order *n*.

For the purpose of this present section let us use the notation

$$A_i B_j = AB = R \quad . \quad . \quad . \quad . \quad . \quad . \qquad (35)$$

to mean adjacent n-rowed matrices within a determinant and *not* the ordinary product of matrices. On shifting i columns of A past B we obtain

$$B_j A_i = BA = (-)^{ij} R. \quad . \quad . \quad . \quad . \qquad (36)$$

In the last section we considered sums of products of two n-rowed determinants. We now extend the results to products of p such determinants

$$(AL),\ (BM),\ (CN)\ .\ .\ .\ .\ .$$

composed of n-rowed matrices

$$A = A_i,\ \ L = L_{n-i},\ \ B = B_j,\ \ M = M_{n-j},\ .\ .\ .\ . \qquad (37)$$

where $i, j, k, \ldots$ are p positive integers each less than n.

We consider the following determinants $\Delta_1, \Delta_2, \ldots, \Delta_p$, of orders $n, 2n, \ldots, pn$ respectively:

$$\Delta_1 = (AL),\ \Delta_2 = \begin{array}{cccc} A & L & B & . \\ A & . & B & M \end{array},\ \Delta_3 = \left|\begin{array}{cccccc} A & L & B & . & C & \\ & & B & M & C & . \\ A & & B & . & C & N \end{array}\right. \qquad (38)$$

where each matrix $A, B, C, \ldots$ is repeated, while each $L, M, N, \ldots$ occurs once. When expanded as in §8, p. 42 by a Laplace development into determinants of order n, these become

$$\Delta_2 = (\dot{A}L)(\dot{B}M),\ \ \Delta_3 = (\dot{A}L)(\dot{B}M)(\dot{C}N),\ \ldots\ . \qquad (39)$$

for again the zero columns prevent such matrices as L, M, N from being permuted. Also by $\text{row}_2 - \text{row}_1$, $\text{row}_3 - \text{row}_1, \ldots$ in (38) we have

$$\Delta_2 = \left|\begin{array}{cccc} A & L & B & . \\ A & . & B & M \end{array}\right| = \left|\begin{array}{cccc} A & L & B & \\ & -L & . & M \end{array}\right|$$

$$\Delta_3 = \left|\begin{array}{cccccc} A & L & B & . & C & . \\ A & . & B & M & C & . \\ A & . & B & . & C & N \end{array}\right. \quad \begin{array}{cccccc} A & L & B & . & C & . \\ & -L & & M & & \\ & -L & & & . & N \end{array} \qquad (40)$$

&c.

Again expanding by a Laplace development Δ_2 leads to results §9, (I), (II), (III), p. 45, already given, while Δ_3, $\Delta_4 \ldots$ give analogous identities which may be stated as follows:

(i) *Let* $\mathrm{i}+\mathrm{j}+\mathrm{k}+\ldots=\mathrm{n}-\mathrm{h}<\mathrm{n}$, *and* $\mathrm{L}_{\mathrm{n}-\mathrm{i}}=\mathrm{L}_{\mathrm{j}}''\,\mathrm{L}_{\mathrm{k}}'''\ldots\mathrm{L}_{\mathrm{h}}'$, *by partitioning the* $\mathrm{n}-\mathrm{i}$ *columns of* $\mathrm{L}_{\mathrm{n}-\mathrm{i}}$ *into its first* j, *next* k, . . . , *and last* h *columns. Then*

$$(\dot{A}\mathrm{L})\,(\dot{B}\mathrm{M})\,(\dot{C}\mathrm{N})\ldots=(\mathrm{ABC}\ldots\dot{\mathrm{L}}')\,(\dot{\mathrm{L}}''\mathrm{M})\,(\dot{\mathrm{L}}'''\mathrm{N})\ldots, \qquad \text{(I)}$$

(ii) *Let* $\mathrm{i}+\mathrm{j}+\mathrm{k}+\ldots=\mathrm{n}$, *and* $\mathrm{L}_{\mathrm{n}-\mathrm{i}}=\mathrm{L}_{\mathrm{j}}''\,\mathrm{L}_{\mathrm{k}}'''\ldots$, *then*

$$(\dot{A}\mathrm{L})\,(\dot{B}\mathrm{M})\,(\dot{C}\mathrm{N})\ldots=(\mathrm{ABC}\ldots)\,(\dot{\mathrm{L}}''\mathrm{M})\,(\dot{\mathrm{L}}'''\mathrm{N})\ldots. \qquad \text{(II)}$$

(iii) *Let* $\mathrm{i}+\mathrm{j}+\mathrm{k}+\ldots>\mathrm{n}$, *then*

$$(\dot{A}\mathrm{L})\,(\dot{B}\mathrm{M})\,(\dot{C}\mathrm{N})\ldots=0. \quad . \quad . \quad . \qquad \text{(III)}$$

The only difficulty here is to settle the sign on the right hand of (I) and (II). The case of Δ_3 for (I) suffices to justify the result. Writing $L=L''L'''L'$, and interchanging the first j columns L'' of L with the j columns of B, and the k columns L''' with C, we have

$$\Delta_3=(-)^{j+k}\begin{vmatrix} A & B & C & L' & L'' & . & L''' & \\ . & . & . & -L' & -L'' & M & -L''' & \\ . & . & . & -L' & -L'' & . & -L''' & N \end{vmatrix}$$

Expanding and shifting all negative signs out of the factor determinants we have

$$\Delta_3=(-)^{2j+2k}\,(ABC\dot{L}')\,(\dot{L}''M)\,(\dot{L}'''N).$$

Similarly for $\Delta_4, \ldots, \Delta_p$.

We can now enunciate these results in the following statement:

A sum of products of p *determinants each with* n *rows, formed by determinantal permutation of* i *columns of the first,* j *columns of the next, and so on, is identically equal to a sum of such products, a single product, or zero, according as* $\mathrm{i}+\mathrm{j}+\ldots<, =, >\mathrm{n}$. *All columns undergoing derangement on one side of the identity are convolved in one factor on the other side, except in the third alternative case.*

Corollary I.—*If each of* i, j, . . . *is unity, the corresponding series on the left of the identity can be written as a* p-*rowed*

compound determinant, *namely one whose elements are determinants.*

Thus, replacing A, B, C by a, b, c, and L, M, N by λ, μ, ν we have

$$(a\lambda)(b\mu)(c\nu) = \begin{matrix} (a\lambda) & (b\lambda) & (c\lambda) \\ (a\mu) & (b\mu) & (c\mu) \\ (a\nu) & (b\nu) & (c\nu) \end{matrix}$$

where each of λ, μ, ν has currency $n - 1$.

It is useful to use Greek letters for these very important matrices involving $n - 1$ columns.

Corollary II.—*If* L, M, N . . . *have certain columns in common, these columns do not undergo derangement.* For the consequent terms would all be zero.

Corollary III.—*Each identity for* n*-rowed determinants can be extended to* n + m*-rowed determinants, simply by affixing a common matrix* Θ *of currency* m *to each of* L, M, N . . . , *and considering all matrices to have* n + m *rows.*

This is the principle of *extensionals.* Incidental examples have already been given (§8, (29), p. 42).

Proof.—Let $(\dot{A}L)\,(\dot{B}M) = (AB\dot{L}')\,(\dot{L}''M)$, $L = L''L'$, be identity (I) for n-rowed determinants. Also let

$$(AL\Theta)\,(BM\Theta)$$

denote a product of two $n + m$-rowed determinants, where A, L, B, M have received m new rows and Θ has $n + m$ rows and m columns. On permuting $\dot{A}$, $\dot{B}$ in this we have an identity

$$(\dot{A}L\Theta)\,(\dot{B}M\Theta) = \Sigma(AB\ldots)\,(\ldots M\Theta).$$

If a column of the Θ from the first factor is displaced, the consequent term vanishes by corollary II. Thus L alone is deranged and

$$(\dot{A}L\Theta)\,(\dot{B}M\Theta) = (AB\dot{L}'\Theta)\,(\dot{L}''M\Theta).$$

Corollary IV.—*Each fundamental identity remains an identity after any further determinantal permutation has been applied to columns of* M, N . . . *and other arbitrary columns, excluding those of* A, B, C, . . . , L′, L″, L‴, . . . *which are already implicitly convolved.*

For if the new operation has t terms, the fundamental identity is true for each such derangement, so that the sum of these t identities is still an identity.

Example.—Since $(\dot{a}ll')(\dot{b}mm') = (ab\dot{l}')(\dot{l}mm')$ we infer that

$$(\dot{a}ll')(\dot{b}mm')(nX) - (\dot{a}ll')(\dot{b}mn)(m'X) = (ab\dot{l}')(\dot{l}mm')(nX) - (ab\dot{l}')(\dot{l}mn)(m'X),$$

where a new operation $\dot{m}', \dot{n}$ of two terms takes effect.

12. Linear Relation between $n+1$ Linear Forms.

The identity (IV) (p. 46) can also be proved as follows:

We form a vanishing determinant of order $n+1$ from n arbitrary rows followed by

$$\text{row}_{n+1} = x_1 \text{ row}_1 + x_2 \text{ row}_2 + \ldots + x_n \text{row}_n.$$

Thus

$$\begin{vmatrix} b_{11} & b_{21} & b_{31} & \ldots & b_{n1} & a_{11} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ b_{1n} & b_{2n} & b_{3n} & \ldots & b_{nn} & a_{1n} \\ b_{1x} & b_{2x} & b_{3x} & \ldots & b_{nx} & a_{1x} \end{vmatrix} = 0,$$

where $b_{ix} = b_{i1}x_1 + b_{i2}x_2 + \ldots + b_{in}x_n$.

Expanding this by the last row, we have

$$(b_2 b_3 \ldots b_n a_1)\, b_{1x} - (b_1 b_3 \ldots b_n a_1)\, b_{2x} + \ldots + (-)^n (b_1 b_2 \ldots b_n)\, a_{1x} = 0. \quad \text{(V)}$$

If in particular $x_1, x_2, \ldots, x_n$ are the n determinants of the matrix D, which consists of $n-1$ columns, the result reverts to (IV). Since the elements a, b, x are all arbitrary, identity (V) gives the important information formally proved in Chapter V, §1, that

Any n + 1 *homogeneous linear forms in* n *variables are necessarily linearly related.*

For $b_{1x}, \ldots, a_{1x}$ are such forms, and (V) is such a linear relation, provided not all the determinants appearing in (V) vanish.

EXAMPLES

1. Prove by this method if $n = 2$, and $a_x = a_1 x_1 + a_2 x_2$, then

$$(bc)\, a_x + (ca)\, b_x + (ab)\, c_x = 0,$$

$$(bc)(ad) + (ca)(bd) + (ab)(cd) = 0.$$

2. If $n = 3$,

$$(bcd)\,a_x + (cad)\,b_x + (abd)\,c_x = (abc)\,d_x,$$
$$(bcd)\,(aef) + (cad)\,(bef) + (abd)\,(cef) = (abc)\,(def).$$

3. Examine the identity (IV) when D is a unit matrix prolonged by zeros: as

$$D = \begin{bmatrix} 1 & . \\ . & 1 \\ . & . \end{bmatrix}, n = 3;\; D = \begin{matrix} 1 & . & . \\ . & 1 & . \\ . & . & 1 \end{matrix}\;, n = 4;\; D = \begin{matrix} 1 & . & . \\ . & 1 & . \\ . & . & 1 \end{matrix}$$

4. Prove if $i = 1, 2, 3$,

$$(bcd)\,a_i + (cad)\,b_i + (abd)\,c_i = (abc)\,d_i,$$

and generalize this result.

5. Prove as a special case of the Sylvester identity (II),

$$(\dot{a}\dot{b}ef)\,(\dot{c}\dot{d})_{12} = (abcd)\,(ef)_{12},\quad (\dot{a}\dot{b}ef)\,(\dot{c}\dot{d})_{ij} = (abcd)\,(ef)_{ij}.$$

6.
$$(\dot{a}\dot{b}pqr)\,(\dot{c}\dot{d})_{12} = (abcd\dot{r})\,(\dot{p}\dot{q})_{12},$$
$$(\dot{a}\dot{b}pqr)\,(\dot{c}\dot{d})_{ij} = (abcd\dot{r})\,(\dot{p}\dot{q})_{ij}.$$

13. Principle of Duality.

There is still something more to be learnt from these fundamental identities. The identity (I) arising from a product of p n-rowed determinants may in fact be looked on as the result of alternative Laplace expansions of the pn-rowed determinant Δ_p. But, as already remarked, this Δ_p gives rise to $p+1$ equivalent expansions, according as we expand Δ_p directly, or after subtracting its first, second ... or pth matrix layer from all the $p-1$ remaining layers. Thus if $p = 3$, $i + j + k < n$, we take

$$\Delta_3 = \begin{matrix} A & L & B & . & C & . \\ A & . & B & M & C & . \\ A & . & B & . & C & N \end{matrix}$$

where the currencies of the six symbols in order are

$$i,\; n-i,\; j,\; n-j,\; k,\; n-k.$$

Hence by shifting columns to bring ABC before LMN, we have

$${}_3 = (-)^{j(n-i)+k(n-i+n-j)} \begin{matrix} A & B & C & L & & \\ A & B & C & & M & \\ A & B & C & & & N \end{matrix} \tag{41}$$

Writing ϵ_3 to denote this power of negative signs, and R for ABC which is of currency $i + j + k = n - h$, we have

$$\Delta_3 = \epsilon_3 \begin{array}{|llll} R & L & & . \\ R & & M & \\ R & & & N \end{array} \qquad (42)$$

and in general

$$\Delta_p = \epsilon_p \begin{array}{llll} R & & & \\ R & & M & \\ R & & & N \end{array} \qquad (43)$$

As in §7, p. 39, we may, but for sign, transform Δ_p so as to bring *any* letter L, M, N, repeated p times over into the first column, with the other p single letters, including R, in any order diagonally. For this reason *we write* Δ_p *as*

$$(R_{n-h} L_{n-i} M_{n-j} N_{n-k} \dots), \quad . \quad . \quad . \quad (44)$$

or shortly $(RLMN \dots)$. It can then be shown that *the inversion law of these symbols is exactly the same as for the* n*-rowed determinant*

$$(E_h A_i B_j C_k \dots) = (EABC \dots)$$

regarded as an expression in the $p + 1$ *matrices of* n *rows,* E, A, B, C, . . . , *of currency indicated by the suffixes.*

Manifestly

$$(E_h A_i B_j C_k \dots) = (-)^{hi} (A_i E_h B \; C_k \dots), \quad . \quad (45)$$

and correlatively it will shortly be proved that

$$(RLMN \dots) = (-)^{hi} (LRMN \dots). \quad . \quad . \quad (46)$$

Sometimes we need to partition one of these matrices, and therefore require a more explicit notation, namely, a full stop between the several matrices. Thus if $R = ABC$,

$$\Delta_3 = (RLMN) = (ABC \, . \, LMN), \quad . \quad . \quad (47)$$

which is also written with a vertical line instead of a stop,

$$(ABC \mid LMN).$$

On reference to (II) we have by expansion

$$(ABC \mid LMN) = (\dot{A}L)(\dot{B}M)(\dot{C}N). \quad . \quad . \quad (48)$$

The currency suffixes are suppressed only when all the matrices have expressly been defined. Otherwise they are necessary.

When several matrices in $(RLMN \ldots)$ are not denoted by single letters, the full stops between are essential. Thus if $a, b, c, \ldots$ are all of unit currency, denoting columns of four-rowed determinants we might have

$$\begin{aligned}(ab \,.\, cde \,.\, fgh) &= (\dot{a}cde)\,(\dot{b}fgh)\\ &= (acde)\,(bfgh) - (bcde)\,(afgh), \quad . \quad (49)\end{aligned}$$

or again

$$(aa'a'' \,.\, bb'b'' \,.\, cc'c'' \,.\, dd'd'') = (\dot{a}bb'b'')\,(\dot{a}'cc'c'')\,(\dot{a}''dd'd''). \quad (50)$$

Proof of the Law of Transposition.

A reference to the original definition of $\Delta_2, \Delta_3, \ldots, \Delta_p$ shows that it is enough to prove the law for adjacent transposition of R, L and of L, M.

Taking the L, M case first and using Δ_3 for brevity, we have

$$\begin{aligned}\Delta_3 = (RLMN) &= (ABC \mid LMN)\\ &= (\dot{A}L)\,(\dot{B}M)\,(\dot{C}N)\\ &= (\dot{B}M)\,(\dot{A}L)\,(\dot{C}N) \quad . \quad (51)\end{aligned}$$

by the commutative law of ordinary multiplication applied to each term of this series. This yields

$$\Delta_3 = (BAC \mid MLN).$$

But by (45) the interchange of AB induces ij changes of sign in R. Hence

$$\Delta_3 = (RLMN) = (-)^{ij}\,(RMLN). \quad . \quad . \quad (52)$$

Next, for the transposition of R, L we have

$$\Delta_3 = (\dot{A}L)\,(\dot{B}M)\,(\dot{C}N) = (ABC\dot{L}')\,(\dot{L}''M)\,(\dot{L}'''N)$$

by the fundamental identity (I), when $L = L''L'''L'$. Here the currency of L' is h $(= n - i - j - k)$, to give n columns for the first factor. Hence by shifting columns

$$\begin{aligned}\Delta_3 &= (-)^{h(n-h)}\,(\dot{L}'ABC)\,(\dot{L}''M)\,(\dot{L}'''N)\\ &= (-)^{h(n-h)}\,(L'L''L''' \mid RMN).\end{aligned}$$

Again, shifting the h columns L' past the $j+k$ columns L'', L''', we have

$$\begin{aligned}\Delta_3 &= (-)^{h(n-h)+h(j+k)}(L''L'''L' \mid RMN)\\ &= (-)^{hi}(LRMN)\end{aligned}$$

by simplifying the sign index. For Δ_p this index is

$$\begin{aligned}&h(n-h)+h(j+k+\ldots)\\ &\qquad = h(n-h)+h(n-h-i) \equiv hi \pmod 2,\end{aligned}$$

leading to the same result. This proves the law.

Definition of Formal Duality:—*Two* n-*rowed matrices of currency* h *and* n — h *respectively are formal duals of each other for the field of order* n.

Formal duality is a relation between numbers of columns: the elements therein may be quite arbitrary.

In the above investigation A is formal dual of L, B of M, &c., and conversely. We extend the definition to include the determinants given in (45) and (46) as formal duals of each other, and accordingly sum this up with a more symmetrical notation as follows:

Let n be partitioned into $p+1$ positive integers i_1, i_2, $\ldots$, i_{p+1}

$$i_1 + i_2 + \ldots + i_{p+1} = n,$$

so that

$$(n-i_1)+(n-i_2)+\ldots+(n-i_{p+1}) = np.$$

Let $A, B, \ldots, K$ be $(p+1)$ matrices of n rows, of currencies i_1, $i_2, \ldots$ respectively, and $L, M, N, \ldots, Q$ be $(p+1)$ such matrices of currencies $n-i_1$, $n-i_2$, $\ldots$. Then the n-rowed determinant

$$(ABC\ldots K)$$

is formal dual of the np-rowed determinant

$$(LMN\ldots Q).$$

More expressly with currency suffixes inserted, these dual determinants are

$$(A_{i_1}B_{i_2}\ldots K_{i_{p+1}}), \quad (L_{n-i_1}M_{n-i_2}\ldots Q_{n-i_{p+1}}).$$

It is also convenient to have a term to describe what is in fact the really important thing about these determinants, the relation between their currencies.

Definition of Characteristic.—*The sets of integers*

$$[i_1, i_2, \ldots, i_{p+1}], \qquad [n - i_1, n - i_2, \ldots, n - i_{p+1}]$$

are the characteristics of the respective determinants.

For example, in results (49) and (50) the characteristics of the left-hand members are (233) and (3333) respectively. Since they are concerned with the field of order 4, their formal duals are (211) and (1111).

These examples, as illustrations of the more general

$$(ABC \mid LMN) = (\dot{A}L)(\dot{B}M)(\dot{C}N)$$

are extremely useful. Taken in conjunction with the fundamental identities, and especially (I), they readily lead to many elegant results, particularly when special values are given to the matrices as in the following examples of compound determinants, which illustrate the characteristic $(n-1, n-1, \ldots, n-1)$, dual of $(11 \ldots 1)$, (cf. §11, Corollary I, p. 48).

Historical Note.—The chief properties of such compounds which have been recorded by Cauchy (1812), Bazin (1854), Sylvester (1841), Whittaker (1915) are still comparatively unknown. Perhaps this is due, especially in the far-reaching cases given by Sylvester, to the rather cumbrous notation originally adopted. It is hoped that the present notation will help the reader to grasp the principles underlying these theorems.

Muir: *History of Determinants,* Vol. I; Vol. II, especially pp. 58–62, 108.
Whittaker: *Proc. Edinburgh Math. Soc.,* **36** (1918), 107–115.
Turnbull: *Proc. London Math. Soc.,* **2, 22** (1923), 503–507.
Ferrar: *Proc. London Math. Soc.,* **2, 23** (1924).

EXAMPLES

1. Prove

$$\begin{array}{ccc} b_1 & l_1 & \\ b_2 & l_2 & \\ b_3 & . & m_1 \end{array} = (bl)(am) - (al)(bm).$$

2. If each letter denotes a column in a three-rowed determinant, write out in full the identity $(\dot{a}pq)(\dot{b}rs)(\dot{c}tu) = (abc)(\dot{p}rs)\dot{q}tu)$. Prove

$$(\dot{a}bc)(\dot{b}ca)(\dot{c}ab) = (abc)^3.$$

3. If $\alpha = pq$, $\beta = rs$, $\gamma = tu$, prove

$$(\alpha\beta\gamma) = (prs)(qtu) - (qrs)(ptu)$$
$$= -(rpq)(stu) + (spq)(rtu).$$

4. Let $\Delta = (abc \, . \, def \, . \, ghk \, . \, lmn)$ denote the formal dual of the four-rowed determinant $(xyzt)$, each letter representing a column of four elements. Prove

$$\Delta = (\dot{a}def)(\dot{b}ghk)(\dot{c}lmn), \; = -(\dot{d}abc)(\dot{e}ghk)(\dot{f}lmn).$$

5. Prove

$$(abc \, . \, bcp \, . \, cqr \, . \, stu) = (abcp)(bcqr)(cstu).$$

6. Generalize result **5** for a product of $n-1$ determinants expressed as a compound determinant.

7.

$$(\dot{a}yz)(\dot{b}zx)(\dot{c}xy) = (abc)(xyz)^2.$$

8.

$$(\dot{a}yzt)(x\dot{b}zt)(xy\dot{c}t)(xyz\dot{d}) = (abcd)(xyzt)^3.$$

9. Generalize 8. This is *Bazin's Theorem* (cf. V, p. 108).

10. If $[ab \, . \, .]$ is the unit matrix, state the results 7 and 8.

Ans. Cauchy's theorem on the adjugate, §7, p. 67.

11.

$$(\dot{a}yz\theta)(\dot{b}zx\theta)(\dot{c}xy\theta) = (abc\theta)(xyz\theta)^2.$$

12.

$$(\dot{a}yzt\theta)(x\dot{b}zt\theta)(xy\dot{c}t\theta)(xyz\dot{d}\theta) = (abcd\theta)(xyzt\theta)^3.$$

13. Generalize 12. This is *Sylvester's Theorem.* It is the extensional of Bazin's Theorem.

14.

$$(bcd \, . \, acd \, . \, xyt \, . \, xyz) = -(abcd)(\dot{c}xyt)(\dot{d}xyz).$$

15.

$$(bcde \, . \, acde \, . \, abde \, . \, zyxu \, . \, xyzt) = -(abcde)^2(\dot{d}xyzu)(\dot{e}xyzt).$$

16. The generalization of 14 and 15 is *Whittaker's Theorem*.

17. If $[abcd \, . \, . \, .]$ is the unit matrix, and $(xyz\theta)$ is an arbitrary determinant with θ denoting all but the first three columns, state the results 11 and 12.

Ans. Jacobi's theorem on the adjugate, p. 79.

CHAPTER IV

MULTIPLICATION OF MATRICES AND DETERMINANTS

1. Fundamental Laws of Algebra.

We now come to the crucial distinction between matrices and ordinary numbers: they do not obey the same law of multiplication. To gain a clear picture of what is here involved we must first recall the four fundamental laws of algebra, on which the whole superstructure is based. These are:

I. The associative law;
II. The distributive law;
III. The commutative law;
IV. The division law.

The first three lead to six primitive facts concerning ordinary addition and multiplication of ordinary numbers:

$$\begin{array}{lll} \text{I} & \text{(i)}\ (a+b)+c=a+(b+c), & \text{(ii)}\ (a\times b)\times c=a\times(b\times c), \\ \text{II} & \text{(i)}\ a\times(b+c)=a\times b+a\times c, & \text{(ii)}\ (a+b)\times c=a\times c+b\times c, \\ \text{III} & \text{(i)}\ a+b=b+a, & \text{(ii)}\ a\times b=b\times a. \end{array}$$

The first pair render the use of brackets unnecessary for continued sums $a+b+c$, and products abc. The fourth law runs as follows:

IV. *If* $\mathrm{a}\times\mathrm{b}=0$, *then either* $\mathrm{a}=0$ *or* $\mathrm{b}=0$.

As Hölder[1] has shown, *any* algebraic identity concerned with real or complex numbers can be deduced from the above laws I, II, III. For instance,

$$(x+y)(x-y)=x^2-y^2$$

is true, but could not be proved merely by the use of the first five: it also requires the law $xy=yx$. On the other hand

---

[1] *Göttinger Nachrichten*, **2** (1889), p. 34.

$(x + y)(a + b) = xa + ya + xb + yb$ does not need the commutative law. Now this sixth law is by far the most interesting of them all. There can be little doubt that whoever first discovered it was led thereto by the orderly arrangement of a things in each of b rows or of b things in each of a rows.

Again, these first six laws are not entirely independent, for II (ii) can be deduced from the others with the help of III (ii), as the reader can quickly prove. But curiously enough the reverse is not the case. The exclusion of the law $a \times b = b \times a$ actually renders the two distributive laws independent.

These laws hold of other classes of things besides numbers a, b, c and other operations besides addition (+) and multiplication (×). Also some but not all of the laws hold of still more things and operations. For example, similar laws operate in an English sentence. The preceding sentence can be regarded as the sum of nine words, or of forty-seven letters, addition having its ordinary meaning. But if we define multiplication to mean the building of a sentence by putting words into their proper order, then such multiplication does not obey the commutative law. The sentences Brutus stabbed Cæsar, and Cæsar stabbed Brutus mean entirely different things. On the other hand the Latin translation of one of these sentences would obey the commutative law.

About ninety years ago two great pioneers in higher algebra, Hamilton[1] and Grassmann[2], initiated schemes of algebra in which the symbols a, b, c did not obey the commutative law of multiplication, although they satisfied the other five I, II, III (i). The accumulated experience of intervening years has amply justified these daring departures from the traditional rule, for it shows conclusively that between the first five laws and the other two there is a profound cleavage. In other words, many algebraic theorems can be proved without recourse to the commutative law of multiplication, thereby opening for algebra entirely new fields, including matrices, vectors, and those latest adumbrations in physics, q-numbers, for the more obvious annexations.

---

[1] *Dublin Transact.*, **17** (1837), 393. *Lectures on Quaternions* (1853).

[2] *Ausdehnungslehre:* Collected Papers, Vol. I (2).

In terms of the first three laws I–III let us speak of

Linear Associative Algebra,

subdivided into

A. Commutative. B. Non-commutative.

Ordinary (scalar) algebra is of type A: that of Hamilton, Grassmann, and the algebra of matrices is of type B. In type B, law III (ii) breaks down. Were it *not* to break down there would be no *raison d'être* of a matrix theory; for every theorem of ordinary algebra would automatically hold for matrices, vectors and the like.

Nevertheless it is a singular fact that the matrix whose very pattern forces on the eye the propriety of the commutative law should be among the first triumphantly to break it!

2. The Law of Multiplication of Matrices.

This law, which Cayley invented and his successors have approved, takes its rise in the theory of linear transformations. Let us consider a simple case with two variables, before and after transformation. If

$$\begin{aligned} x_1 &= a_1 y_1 + b_1 y_2, & y_1 &= p_1 z_1 + p_2 z_2, \\ x_2 &= a_2 y_1 + b_2 y_2, & y_2 &= q_1 z_1 + q_2 z_2, \end{aligned}$$

we have what is called a system of two linear equations transforming x_1, x_2 to y_1, y_2 and again two equations transforming y_1, y_2 to z_1, z_2. We write

$$X = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad A = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}, \quad Y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}, \quad B = \begin{bmatrix} p_1 & p_2 \\ q_1 & q_2 \end{bmatrix}, \quad Z = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}.$$

A and B are called the matrices of the respective transformations: X, Y, Z are the matrices of the variables. The two transformations are sometimes written as

$$X \rightarrow Y, \quad Y \rightarrow Z.$$

But by eliminating y_1, y_2 we have at once

$$\begin{aligned} x_1 &= (a_1 p_1 + b_1 q_1) z_1 + (a_1 p_2 + b_1 q_2) z_2 \\ x_2 &= (a_2 p_1 + b_2 q_1) z_1 + (a_2 p_2 + b_2 q_2) z_2, \end{aligned}$$

which is manifestly a linear transformation from x_1, x_2 direct to z_1, z_2. If C denote its coefficient matrix, then

$$C = \begin{bmatrix} a_1 p_1 + b_1 q_1, & a_1 p_2 + b_1 q_2 \\ a_2 p_1 + b_2 q_1, & a_2 p_2 + b_2 q_2 \end{bmatrix}.$$

Here we have the suggestion of a product of matrices. In fact C *is defined to be the product of* A *and* B, *namely*

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \times \begin{bmatrix} p_1 & p_2 \\ q_1 & q_2 \end{bmatrix} = \begin{bmatrix} a_1 p_1 + b_1 q_1, & a_1 p_2 + b_1 q_2 \\ a_2 p_1 + b_2 q_1, & a_2 p_2 + b_2 q_2 \end{bmatrix}.$$

In short $A \times B = AB = C$.

Thus the product of matrices is based on that of linear transformations, and we might, for example, indicate this by

$$(X \rightarrow Y)\,(Y \rightarrow Z) = (X \rightarrow Z),$$

meaning, the transformation from X to Y followed by that from Y to Z produces the direct transformation from X to Z. And there is no harm even in speaking of the transformation A, meaning that from X to Y whose coefficient matrix is A.

Now it is at once apparent from the definition that the product BA in general means something different from AB. Thus

$$BA = \begin{bmatrix} p_1 & p_2 \\ q_1 & q_2 \end{bmatrix} \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} = \begin{bmatrix} a_1 p_1 + a_2 p_2, & b_1 p_1 + b_2 p_2 \\ a_1 q_1 + a_2 q_2, & b_1 q_1 + b_2 q_2 \end{bmatrix} = D,$$

which is only equal to AB if all four corresponding elements in D equal those in C: e.g. $a_1 p_1 + b_1 q_1 = a_1 p_1 + a_2 p_2$. Obviously this is not true in general since the eight elements $a_1 \ldots q_2$ are arbitrary numbers.

Next, we extend this definition to include other than square matrices, by analogy from the case

$$\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} y_1 & 0 \\ y_2 & 0 \end{bmatrix} = \begin{bmatrix} a_1 y_1 + b_1 y_2, & 0 \\ a_2 y_1 + b_2 y_2, & 0 \end{bmatrix}.$$

We *define*

$$AY = \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} a_1 y_1 + b_1 y_2 \\ a_2 y_1 + b_2 y_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = X.$$

Hence the formula $X = AY$ literally represents the linear transformation of a column of variables x_1, x_2 to that of y_1, y_2. Similarly

$$Y = BZ, \quad X = CZ, \quad X = (AB)Z, \quad X = A(BZ).$$

This last suggests that for any conformable square matrices the associative law holds, which is in fact the case; namely

$$P(QR) = (PQ)R,$$

which will be proved in §4, p. 63.

EXAMPLES

1. If $I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, and A, B, C are arbitrary square matrices of four elements each, prove

(i) $IA = AI = A$.
(ii) $OA = AO = O$,
(iii) $A(B + C) = AB + AC$,
(iv) $(A + B)C = AB + BC$.

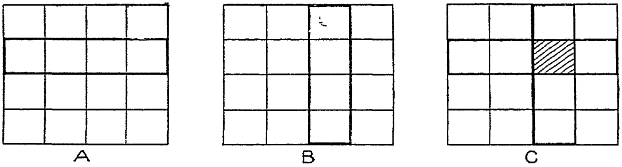
2. If x, y are scalar numbers, $A(xI) = xAI = xA$. Also $A(xB) = (xA)B$, $xAyB = yxAB$.

3. Product of Square Matrices of Order n.

The general case is now straightforward. Let A, B be two such matrices of order n,

$$A = \begin{bmatrix} a_1 & b_1 & \dots & m_1 \\ \cdot & \cdot & \cdot & \cdot \\ a_n & b_n & \dots & m_n \end{bmatrix}, \qquad B = \begin{bmatrix} p_1 & p_2 & \dots & p_n \\ q_1 & q_2 & \dots & q_n \\ \cdot & \cdot & \cdot & \cdot \\ t_1 & t_2 & \dots & t_n \end{bmatrix}$$

Let the ith row of A be denoted by $r_1, r_2, \dots, r_n$ and the jth column of B by $c_1, c_2, \dots, c_n$. We form a new matrix C by



weaving this ith row of A into the jth column of B by the following rule which defines what is called the *inner product* $(r \mid c)$ of these two sets of n elements

$$(r \mid c) = r_1c_1 + r_2c_2 + \dots + r_nc_n.$$

This sum is taken for the (i, j)th element of C; and by choosing all values 1, 2, ..., n for i and for j we completely determine the n^2 elements of C. Evidently this rule includes the case already cited when $n = 2$. Also the notation $(r \mid c)$ is entirely in agreement with that of §13 (48), p. 52, as will be apparent later in Chap. V, p. 81.

This notation $(r \mid c)$, sometimes written r_c, for the inner product is admirable for exhibiting the product of two matrices. For instance,

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \times \begin{bmatrix} p_1 & q_1 & r_1 \\ p_2 & q_2 & r_2 \\ p_3 & q_3 & r_3 \end{bmatrix} = \begin{bmatrix} (a \mid p) & (a \mid q) & (a \mid r) \\ (b \mid p) & (b \mid q) & (b \mid r) \\ (c \mid p) & (c \mid q) & (c \mid r) \end{bmatrix}$$

Product of Rectangular Matrices.

The definition may easily be extended to the product of rectangular matrices. It is only necessary for the fore and after factors to have an equal length of row and column respectively. Thus for a length of three terms

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \times \begin{bmatrix} p_1 & q_1 \\ p_2 & q_2 \\ p_3 & q_3 \end{bmatrix} \quad \begin{bmatrix} (a \mid p) & (a \mid q) \\ (b \mid p) & (b \mid q) \end{bmatrix},$$

and the extreme instance is the single row with the single column

$$[r_1 \quad r_2 \qquad r_n] \times \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = r_1c_1 + \dots + r_nc_n = (r \mid c).$$

EXAMPLES

$$\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \\ -1 & 0 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 2 & 0 & 1 \end{bmatrix}.$$

2. Reverse the order in the above.

3. Form the square and the cube of the matrix $A = \begin{bmatrix} -1 & 1 & -1 & 1 \\ -3 & 2 & -1 & \\ -3 & 1 & . & \\ -1 & . & . & . \end{bmatrix}$

Ans. A^3 is the *unit* matrix (§8).

4. Prove that if $\{m, r\}$ denote an $m \times r$ matrix, $\{m, r\} \times \{r, n\} = \{m, n\}$.

5. If X denote a column of n variables $x_1, x_2, \ldots, x_n$, and if Y, Z have similar meanings for y_i, z_i, generalize the results of linear transformation already given when $n = 2$. Namely, $X = AY$, $Y = BZ$, $X = CZ = ABZ$.

6. The single column c premultiplying the single row r, each of n elements, gives the $n \times n$ matrix $[c_i r_j]$ which is of unit rank.

4. Double Suffix Notation of Multiplication.

It is here that the double suffix notation also is very useful. Let

$$A = [a_{ij}], \quad B = [b_{ij}],$$

be two square matrices of order n, so that a_{ij} is the element in the ith row and jth column. Then the inner product of the ith row of A and kth column of B is c_{ik} if

$$\begin{aligned} c_{ik} &= a_{i1} b_{1k} + a_{i2} b_{2k} + \ldots + a_{in} b_{nk} \\ &= \sum_j a_{ij} b_{jk} \\ &= (a_i \mid b_k). \end{aligned}$$

Thus if $C = AB$, then

$$c_{ik} = \Sigma\, a_{ij} b_{jk} = (a_i \mid b_k). \quad . \quad . \quad . \quad . \quad (1)$$

The notation lends itself to continued products. So we should have

$$x_{il} = \sum_j \sum_k a_{ij} b_{jk} c_{kl}$$

in the case when

$$X = ABC. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

It will be seen that the suffixes on the right are linked in pairs, with unpaired end suffixes i, l answering to those of x_{il} on the left. The double series has n^2 terms; and since we obtain exactly the same result whether we sum for k first and for j next, or vice versa, we are justified in dropping the brackets in $A(BC)$ and $(AB)C$. They mean the same thing.

5. The Division Law.

The law IV (p. 57), that if $xy = 0$ then either x or y must be zero does not hold for matrices. For example, let

$$A = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} b & b \\ -a & -a \end{bmatrix},$$

then

$$AB = \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} \begin{bmatrix} b & b \\ -a & -a \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0.$$

Yet neither factor A nor B is the zero or null matrix. Consequently this law fails for A and B.

At the same time the law holds for certain matrices; for instance, if the determinant $|B|$ were non-zero it could be proved that $AB = 0$ only if $A = 0$.

The reason why this is called the *division* law rests on the fact that it defines a unique process converse to multiplication. For suppose $AB = C$ and $AD = C$ are two instances of multiplication in linear associative algebra which give the same product C. Then $A(B - D) = AB - AD = C - C = 0$; so that if $A \neq 0$ the law shows that $B - D = 0$ or $B = D$. Hence there is only one possible factor following a given A in a product AB which can produce a given C.

In this case A is called the left or fore factor and B the right or after factor of C. Looking at the same relation conversely we consider B to be the quotient when C is divided by A; more precisely, B is the quotient of left or fore division of C by A.

Similarly, if both $AB = C$, $EB = C$ it follows from the same law that $A = E$. Hence there is only one factor A preceding a given B in a product AB which can produce a given C. So conversely A is the quotient of right or after division of C by B.

To sum up, there are in the case where AB, BA are not necessarily equal, *two* sorts of multiplication—fore and after multiplication, and *two* analogous sorts of division. But these last require the division law to hold. Algebra for which this law is true is called *division algebra*.[1]

---

[1] The reader who wishes to study this very important law should consult Dickson, *Linear Algebras* (Cambridge, 1914), particularly the theorem of p. 10, due to Frobenius, *Crelle*, **84** (1878), p. 59; also Dickson, *Algebras and their Arithmetics* (Chicago, 1923) and the revised and German edition, *Algebren und ihre Zahlentheorie* (Zürich, 1927).

EXAMPLES

1. A matrix product with the zero matrix for factor is zero.

[Note that *two* proofs are necessary, for left and right factor.]

2. If $ad \neq bc$, $xt \neq yz$, then the product $\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x & y \\ z & t \end{bmatrix}$ cannot be zero.

3. The first five laws always hold and the sixth sometimes holds for matrices.

Consider $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$, $B = \begin{bmatrix} 4 & -2 \\ -3 & 1 \end{bmatrix}$, $AB = BA = -2I$.

We consider this sixth law further in §§8 and 9 below.

6. Products of Determinants.

The product of two determinants each of order n *is itself a determinant of the same order.*

Consider, for example, the product of $(abc)(\alpha\beta\gamma)$. It can be written as a six-rowed determinant:

$$\Delta = \begin{vmatrix} a_1 & b_1 & c_1 & -1 & \cdot & \cdot \\ a_2 & b_2 & c_2 & \cdot & -1 & \cdot \\ a_3 & b_3 & c_3 & \cdot & \cdot & -1 \\ \cdot & \cdot & \cdot & \alpha_1 & \alpha_2 & \alpha_3 \\ \cdot & \cdot & \cdot & \beta_1 & \beta_2 & \beta_3 \\ \cdot & \cdot & \cdot & \gamma_1 & \gamma_2 & \gamma_3 \end{vmatrix}$$

since the $\left(\begin{array}{c} 3 \\ \hdashline 3 \end{array}\right)$ Laplace development contains only the one term

$$(abc)(\alpha\beta\gamma).$$

Replace in Δ rows four, five, and six by the following equivalents:

$$\begin{aligned} &\mathrm{row}_4 + \alpha_1\ \mathrm{row}_1 + \alpha_2\ \mathrm{row}_2 + \alpha_3\ \mathrm{row}_3, \\ &\mathrm{row}_5 + \beta_1\ \mathrm{row}_1 + \beta_2\ \mathrm{row}_2 + \beta_3\ \mathrm{row}_3, \\ &\mathrm{row}_6 + \gamma_1\ \mathrm{row}_1 + \gamma_2\ \mathrm{row}_2 + \gamma_3\ \mathrm{row}_3. \end{aligned}$$

Hence $\Delta =$

$$\begin{array}{llllll} a_1 & b_1 & c_1 & -1 & \cdot & \\ a_2 & b_2 & c_2 & \cdot & -1 & \\ a_3 & b_3 & c_3 & & & -1 \\ a_1\alpha_1+a_2\alpha_2+a_3\alpha_3 & b_1\alpha_1+b_2\alpha_2+b_3\alpha_3 & c_1\alpha_1+c_2\alpha_2+c_3\alpha_3 & & & \\ a_1\beta_1+a_2\beta_2+a_3\beta_3 & b_1\beta_1+b_2\beta_2+b_3\beta_3 & c_1\beta_1+c_2\beta_2+c_3\beta_2 & & & \\ a_1\gamma_1+a_2\gamma_2+a_3\gamma_3 & b_1\gamma_1+b_2\gamma_2+b_3\gamma_3 & c_1\gamma_1+c_2\gamma_2+c_3\gamma_3 & & & \end{array}$$

Again expand by the $\begin{pmatrix}3\\ \cdots\\ 3\end{pmatrix}$ development and we obtain

$$\begin{vmatrix} a_1\, b_1\, c_1 \\ a_2\, b_2\, c_2 \\ a_3\, b_3\, c_3 \end{vmatrix} \times \begin{vmatrix} \alpha_1\, \alpha_2\, \alpha_3 \\ \beta_1\, \beta_2\, \beta_3 \\ \gamma_1\, \gamma_2\, \gamma_3 \end{vmatrix} = (-)^3 \begin{vmatrix} -1 & . & . \\ . & -1 & . \\ . & . & -1 \end{vmatrix} \begin{vmatrix} (a|\alpha) & (b|\alpha) & (c|\alpha) \\ (a|\beta) & (b|\beta) & (c|\beta) \\ (a|\gamma) & (b|\gamma) & (c|\gamma) \end{vmatrix},$$

The final determinant is called the product by *columns and rows*.

It is often written with a_α replacing the notation $(a \mid \alpha)$ for $\Sigma a_\iota \alpha_i$, which gives the following product by *rows and columns*:

$$(\alpha\beta\gamma)(abc) = \begin{vmatrix} a_\alpha & b_\alpha & c_\alpha \\ a_\beta & b_\beta & c_\beta \\ a_\gamma & b_\gamma & c_\gamma \end{vmatrix} \qquad (3)$$

Corollary.—The product rule of matrices agrees with that of determinants. Thus if

$$AB = C, \text{ then } |A|\,|B| = |C|.$$

The converse is not true; for owing to the great variety of patterns of the same determinant

$$\Delta = \Sigma \pm a_1 b_2 \ldots m_n$$

made by interchanging rows and columns without disturbing the actual contents of a row or column, there are many equivalent product determinants. This is by no means true of matrices. Herein is the manifest difference between matrices and determinants, as interchange of rows and columns gives a different multiplication rule valid for determinants but not for matrices. But undoubtedly the rule as given in (3), of weaving columns into rows, thus forming all possible inner products of a row of the first factor and a column of the second factor, is the best to store in the memory. But occasionally it is useful to multiply *rows* by *rows*, or *columns* by *columns*.

7. Reciprocal and Adjugate Determinants.

Let us now use capital letters to denote co-factors of elements in a determinant

$$\Delta = | a_1 b_2 c_3 \ldots m_n |,$$

so that

$$\Delta = a_1 A_1 + a_2 A_2 + \ldots + a_n A_n = a_A$$

in the notation for an inner product just explained.

Since $| b_1 b_2 c_3 \ldots m_n |$ has two equal columns, it vanishes. Thus

$$0 = b_1 A_1 + b_2 A_2 + \ldots + b_n A_n,$$

which may be written $\quad b_A = 0.$

In general $\quad p_Q = 0, \quad p_P = \Delta,$

for exactly the same reasons. Correlatively we have

$$\Delta = a_1 A_1 + b_1 B_1 + \ldots + m_1 M_1,$$

which we typify by 1_1, and

$$0 = a_2 A_1 + b_2 B_1 + \ldots + m_2 M_1,$$

which we typify by 2_1. So in general

$$\begin{array}{ll} i_j = 0 & i \neq j \\ i_i = \Delta & i = 1, 2, \ldots\ n. \end{array}$$

Definition.—*The determinant*

$$\bar{\Delta} = \begin{array}{ccccc} A_1 & B_1 & C_1 & \ldots & M_1 \\ A_2 & B_2 & C_2 & & M_2 \\ & & & & \\ A_n & B_n & C_n & \ldots & M_n \end{array}$$

is called the adjugate of Δ. Its elements are the co-factors of the corresponding elements of Δ.

Since by multiplication

$$\Delta \bar{\Delta} = \begin{vmatrix} a_A & b_A & \ldots & m_A \\ a_B & b_B & \ldots & m_B \\ \cdot & \cdot & \cdot & \cdot \\ a_M & b_M & \ldots & m_M \end{vmatrix} = \begin{vmatrix} \Delta & 0 & 0 & \ldots & 0 \\ 0 & \Delta & 0 & \ldots & 0 \\ 0 & 0 & \Delta & \ldots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \ldots & \Delta \end{vmatrix} = \Delta^n$$

it follows that provided $\Delta \neq 0$,

$$\bar{\Delta} = \Delta^{n-1}$$

a very beautiful property, due to Cauchy.[1]

Definition.—The determinant whose elements are those of the adjugate each divided by Δ is the *reciprocal* of Δ.

---

[1] *Journal de l'École Polytechnique*, **17** (1815), 82. The same property was proved in Ex. 10, p. 56, without assuming Δ to be nonzero.

It can be written as Δ^{-1} or $\frac{1}{\Delta}$, since

$$\begin{vmatrix} \frac{A_1}{\Delta}, & \frac{B_1}{\Delta} & \cdots \\ \frac{A_2}{\Delta}, & \frac{B_2}{\Delta} & \cdots \end{vmatrix} = \Delta^{-n} \begin{vmatrix} A_1 & B_1 & \cdots \\ A_2 & B_2 & \cdots \\ \cdot & \cdot & \cdot \end{vmatrix} = \Delta^{-n}\bar{\Delta} = \frac{1}{\Delta}.$$

We have multiplied each row of the determinant by Δ, at the same time dividing the result by Δ^n. This is an example where the matrix theory would differ. Thus by matrix definition

$$\begin{bmatrix} \frac{A_1}{\Delta}, & \frac{B_1}{\Delta} \\ A_2 & \frac{B_2}{\Delta} \end{bmatrix} = \frac{1}{\Delta}\begin{bmatrix} A_1 & B_1 \\ A_2 & B_2 \end{bmatrix}$$

where Δ^{-1} is now the factor instead of Δ^{-n}.

8. The Index Law and the Reversal Law of a Matrix.

There is no ambiguity in writing A^2 for the product of a square matrix A with itself. It follows by the associative law that if r is a positive integer, A^r is a useful abbreviation for the continued product of r equal matrices A. Further, as in ordinary algebra, the index law

$$A^r A^s = A^{r+s}$$

will hold for all positive integral values of r and s. And, if $|A| \neq 0$, we may allow r, s to be any integers, zero and negative included, by adding the following two definitions, for the *unit matrix* and the *reciprocal matrix.*

Definition of Unit Matrix.—*The square matrix whose leading diagonal elements are each unity, all other elements being zero, is called the unit matrix.*

Definition of Reciprocal Matrix.—*The square matrix* $[\alpha_{ij}]$ *whose* (i, j)*th element is the co-factor of* a_{ji} *in the determinant* $|a_{ji}|$, *divided by the determinant itself, is called the reciprocal of the matrix* $[a_{ji}]$.

The reasons for these definitions are simple, for they are analogous to those of elementary algebra. Thus, if I is the unit matrix and M is a square matrix of the same order n, then, taking $n = 3$,

$$I = \begin{bmatrix} 1 & . & . \\ . & 1 & . \\ . & . & 1 \end{bmatrix}, \quad M = \begin{bmatrix} a_1 & b_1 & c_1 \\ & & \\ a_3 & b_3 & c_3 \end{bmatrix} \tag{4}$$

whence by actual multiplication

$$IM = M = MI,$$

so that the unit matrix as a factor leaves another matrix unchanged. Indeed the unit matrix is seen to be commutative with another matrix.

Also if N is the reciprocal of M, then by definition

$$N = \begin{bmatrix} \frac{A_1}{\Delta} & \frac{A_2}{\Delta} & \frac{A_3}{\Delta} \\ \frac{B_1}{\Delta} & \frac{B_2}{\Delta} & \frac{B_3}{\Delta} \\ \frac{C_1}{\Delta} & \frac{C_2}{\Delta} & \frac{C_3}{\Delta} \end{bmatrix} \tag{5}$$

where A_1 is co-factor of a_1 in Δ, the determinant of M. Thus by actual multiplication

$$NM = \begin{matrix} 1 & . & \\ & 1 & \\ & & 1 \end{matrix} = I. \tag{6}$$

Similarly $MN = I$. Thus *a matrix is commutative with its reciprocal*, and if we define M^{-1} to mean the reciprocal of M we have the result

$$MM^{-1} = I = M^{-1}M. \tag{7}$$

This allows us to use the notation M^{-r} to mean indifferently the rth power of the reciprocal of M or the reciprocal of the rth power of M, provided $|M| \neq 0$. Also we have $M^0 = I$.

As examples of this index law the reader should prove the following results, noticing the curious feature which emerges—one of great importance in all non-commutative algebra—that *the inverse or reciprocal of a product of factors reverses the order of the factors.*

$$\left.\begin{aligned}(AB)^{-1} &= B^{-1}A^{-1}\\ (ABC\ldots K)^{-1} &= K^{-1}\ldots C^{-1}B^{-1}A^{-1}\\ (A^{-1}B^{-1})^{-1} &= BA.\end{aligned}\right\} \quad . \quad (8)$$

Further, *the same reversal is true of the operation of transposing a matrix,* which is denoted by an accent.

$$(AB)' = B'A', \text{ \&c.} \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Lastly, *the operations of inversion and transposition are commutative*:

$$(A')^{-1} = (A^{-1})'. \quad . \quad . \quad . \quad . \quad . \quad . \quad (10)$$

Singular Matrices.

When $|A|$ vanishes, the matrix A is said to be singular. It has no reciprocal, although there still exists the matrix of elements $[\alpha_{ji}]$ where α_{ij} is the co-factor of a_{ij} in $|A|$, which has certain properties analogous to those of a reciprocal matrix. This is called the *adjugate matrix.*

It is easy to adapt the result (6). Thus the product of A and its adjugate gives

$$[\alpha_{ji}] = \begin{vmatrix} |A| & \cdot & \cdot \\ \cdot & |A| & \cdot \\ \cdot & \cdot & |A| \end{vmatrix} = |A|\,I = |A \qquad (11)$$

This is also true for n rows and columns. In particular *the product of a singular matrix and its adjugate is the zero matrix.*

Here is an example where the division law fails: neither factor need necessarily be zero.

9. Summary of Laws of Matrices.

We conclude this chapter by summarizing our results, for we now hold all the fundamental laws of matrix algebra. First there are three fundamental operations,

Addition,
Multiplication,
Transposition,

this last being new, for it does not occur in ordinary algebra.

Denoting transposition by an accent, the laws which govern matrices are

$$(A+B)+C=A+(B+C) \qquad (A\times B)\times C=A\ \times(B\times C)$$
$$A\times(B+C)=AB+AC \qquad (A+B)\times C=AC+BC$$
$$A+B=B+A$$

$$(A+B)'=A'+B', \quad (A')'=A, \quad (AB)'=B'A',$$
$$(A^{-1})^{-1}=A, \quad (AB)^{-1}=B^{-1}A^{-1},$$
$$(A^{-1})'=(A')^{-1}.$$

The commutative law of multiplication fails in general, but holds at any rate in certain cases, namely

AX *and* XA *are equal when* X *is zero, the unit matrix, a power of* A, *or a scalar matrix.*

The *scalar matrix*
$$\begin{bmatrix} p & . & . & . & . \\ & p & & & \\ & & & & \\ & & & & p \end{bmatrix} = pI,$$

is a device for expressing an ordinary number p as a matrix. In this way, ordinary algebra can be thought of as a particular case of matrix algebra—when all matrices involved are unit matrices all of the same order. Thus, for instance, $p^2-q^2=(p+q)\ (p-q)$ in matrix algebra would be $P^2-Q^2=(P+Q)\,(P-Q)$ where $P=pI$, $Q=qI$. Also $pA=pIA=A(pI)=Ap$, a result which incorporates and extends *scalar* multiplication as defined on p. 34.

In the following examples capital letters mean matrices of order n, small letters mean ordinary numbers.

EXAMPLES

1. Why is $A^2-B^2\neq(A+B)(A-B)$ in general?

2. Prove $A^2-I^2=(A+I)(A-I)=(A-I)(A+I)$.

3. $A^2-(\lambda+\mu)A+\lambda\mu I=(A-\lambda I)(A-\mu I)$.

4. If A, B are two-rowed matrices, and B commutes with A then $B=\lambda A+\mu I$.

5. If $f(A)=p_0A^q+p_1A^{q-1}+\ldots+p_iA^{q-i}+\ldots+p_nI$, where $p_0,\ldots,p_n$ are scalar and q is a positive integer, prove $f(A)$ is a matrix of order n which commutes with A.

6. If $g(A)$ is another such polynomial, prove that $\{g(A)\}^{-1}$ is another such matrix which commutes with A, provided the determinant $|g(A)|$ does not vanish.

7. Prove $f(A) \times (g(A))^{-1} = (g(A))^{-1} \times f(A)$. Why is the notation $\frac{A}{B}$ ambiguous, but $\frac{f(A)}{g(A)}$ not so?

8. If $|B| \neq 0$, and $AB = BA$, prove that $AB^{-1} = B^{-1}A$.

9. If $A = \begin{bmatrix} . & -\tan\frac{\alpha}{2} \\ \tan\frac{\alpha}{2} & . \end{bmatrix}$ prove $\frac{I+A}{I-A} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$.

10. Any two rational functions $\varphi(A)$, $\psi(A)$ of a single matrix A commute with one another, and hence they only differ in behaviour from scalar numbers in failing to obey the division law.

11. Prove $(I - AB)A(I + BA) = (I + AB)A(I - BA)$.

12. Prove the product $\begin{bmatrix} 0 & c & -b \\ -c & 0 & a \\ b & -a & 0 \end{bmatrix} \begin{bmatrix} a^2 & ab & ac \\ ab & b^2 & bc \\ ac & bc & c^2 \end{bmatrix}$ is zero.

13. If $AB = 0$ it does not follow that $BA = 0$.

$$\text{Try} \begin{bmatrix} a & b \\ . & . \end{bmatrix} \begin{bmatrix} b & . \\ -a & . \end{bmatrix}.$$

14. If $X = [x_{ij}]$, $A = [a_{ij}]$, $B = [b_{ij}]$, ... denote square matrices of order n, prove that $\sum_i x_{ii}$ denotes the sum of the elements in the leading diagonal of X.

If s_X denote this sum, prove

$$s_{AB} = \sum_i \sum_j a_{ij} b_{ji}, \quad s_{ABC} = \sum_i \sum_j \sum_k a_{ij} b_{jk} c_{ki},$$

each summation running 1, 2, ..., n.

15. Prove $s_{AB} = s_{BA}$, $s_{ABC} = s_{BCA} = s_{CAB}$.

16. The sum of the elements in the leading diagonal is the same for all matrices X, Y, ... where

$$X = ABC \ldots H, \quad Y = BC \ldots HA, \ldots.$$

obtained by cyclic symmetry.

17. Prove $s_X = s_{X'}$.

18. Prove that, if $n = 2$, $\begin{bmatrix} \frac{\partial y}{\partial x_{11}} & \frac{\partial y}{\partial x_{21}} \\ \frac{\partial y}{\partial x_{12}} & \frac{\partial y}{\partial x_{22}} \end{bmatrix} = 2X$, where $y = s_X{}^2$.

19. Denoting this matrix in general by $\left[\frac{\partial y}{\partial x_{ji}}\right]$ prove

$$\left[\frac{\partial y}{\partial x_{ji}}\right] = rX^{r-1} \text{ when } y = s_X{}^r.$$

CHAPTER V

LINEAR EQUATIONS. THE THEOREM OF CORRESPONDING MATRICES. FURTHER THEOREMS

1. Matrices and Linear Equations. Rank.

We now replace the definition of rank given in §5, p. 10, by a more practical statement.

Definition of Rank.—*A matrix of* m *rows and* n *columns has rank* r, *when not all its minor determinants of order* r *vanish, while all of order* r + 1 *do so.*

It follows at once by a Laplace development that not all minors of order less than r vanish, whereas all of order greater than r do so. Also $r \leqslant n$, $r \leqslant m$.

Again, since non-zero determinants exist for all values of n, the unit determinant $|\delta_{ij}|$ being such an one, a matrix *can* have rank equal to the smaller of n or m.

When $r = n = m$ the matrix is *non-singular*, and for the subsequent theory of invariants this is by far the most important case. When $r < n$ or $r < m$ the matrix is *singular*.

It can now be shown that the new definition implies the earlier property, that a matrix of rank r has r linearly independent rows and also r such columns, but not $r + 1$ such rows or $r + 1$ such columns.

Proof.—At least one set of r columns exists between which a relation

$$\mu_1 \operatorname{col}_a + \mu_2 \operatorname{col}_b + \ldots + \mu_r \operatorname{col}_g = 0 \quad . \quad . \quad (1)$$

is impossible. For let the r columns $a, b, \ldots, g$ belong to any one r-rowed minor $\Delta_r \equiv |a_i b_j \ldots g_h|$ which does not vanish. Then the assumed relation (1) requires among others, the r equations

$$\begin{matrix} \mu_1 a_i + \mu_2 b_i + \ldots + \mu_r g_i = 0 \\ . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \\ \mu_1 a_h + \mu_2 b_h + \ldots + \mu_r g_h = 0 \end{matrix} \qquad (2)$$

Multiplying these respectively by the co-factors of $a_i, \ldots, a_h$ in Δ_r, and adding, we obtain

$$\mu_1 \,|\, a_i b_j \ldots g_h \,| + \mu_2 \,|\, b_i b_j \ldots g_h \,| + \ldots = 0, \text{ i.e. } \mu_1 \,.\, \Delta_r = 0, \quad (3)$$

whence $\mu_1 = 0$, since $\Delta_r \neq 0$. Similarly each μ vanishes and no relation (1) exists. The corresponding case of rows can be treated in the same way.

Now let $q_i, q_j, \ldots, q_k$ be $r + 1$ arbitrary non-zero numbers, and $[a_i b_j \ldots p_k]$ be any minor matrix of order $r + 1$ of a given matrix M, formed by attaching row k and column p to those represented in Δ_r. Then, from the fundamental identity (IV) (cf. §12, Ex. 3, 4, p. 51) we deduce the further identity

$$\begin{aligned}(qb \ldots gp)_{ij\ldots k} a_i + (aq \ldots gp)_{ij\ldots k} b_i + \ldots + (ab \ldots gq)_{ij\ldots k} p_i \\ = (ab \ldots gp)_{ij\ldots k} q_i. \quad . \quad . \quad . \quad (4)\end{aligned}$$

Since M has rank r, this minor determinant $(ab \ldots p)_{ij\ldots k}$ of order $r + 1$ vanishes. Hence the left-hand sum also is identically zero. Thereby the $r + 1$ elements $a_i, b_i \ldots p_i$ of row i in M are linearly related: namely

$$\lambda_1 a_i + \lambda_2 b_i + \ldots + \lambda_{r+1} p_i = 0, \quad . \quad . \quad . \quad (5)$$

where $\lambda_1 = (qb \ldots p)_{ij\ldots k}$, &c. Here λ_{r+1} at least cannot vanish identically because, in it, Δ_r is co-factor of q_k. Similarly the same relation holds for rows $j, \ldots, k$. Further it holds for any other row l, since

$$\lambda_1 a_l + \lambda_2 b_l + \ldots + \lambda_{r+1} p_l = (aqb \ldots p)_{lij\ldots k}, \quad (6)$$

as expansion by row l of this determinant shows, the term involving q_l being zero. But this determinant vanishes on expansion by its column q, since each co-factor so formed is a determinant of order $r + 1$ belonging to M. Combining these results, the $r + 1$ *columns* are linearly related; namely

$$\lambda_1 \operatorname{col}_a + \lambda_2 \operatorname{col}_b + \ldots + \lambda_{r+1} \operatorname{col}_p = 0. \quad . \quad . \quad (7)$$

This expresses any column p in terms of r suitably chosen columns. Similarly we prove any $r + 1$ rows to be linearly related. This proves the theorem.

Corollary.—*A matrix and its transposed have the same rank.*

2. Application to Linear Equations.

A matrix M of n rows and m columns is closely connected with the theory of n homogeneous equations linear in m variables, or m such equations in n variables.

Thus if $a_x = a_1 x_1 + a_2 x_2 + \dots + a_n x_n$ with like abbreviations $b_x, c_x, \dots, k_x$, the matrix

$$M' = \begin{bmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & & b_n \\ & & & \\ k_1 & k_2 & & \end{bmatrix}$$

is that of the m linear equations

$$a_x = 0, \quad b_x = 0, \qquad k_x = 0.$$

To say $\lambda \text{ row}_a + \mu \text{ row}_b + \dots = 0$ is in effect the same as to say

$$\lambda a_x + \mu b_x + \dots \equiv 0 \; \{x\},$$

namely that this is an identity for all values of x. Hence the rank r of M' determines exactly how many of the equations are effectively independent. So among $a_x, b_x, \dots, k_x$ exactly r independent forms can be chosen. Let them be the first r:—$a_x, b_x, \dots, g_x$. Also among the columns exactly r are independent. Let them be columns $1, 2, \dots, r$. As we are dealing with equations, there is no loss of generality in making these assumptions.

If now $r = n \leqslant m$, we have n equations $a_x = 0, b_x = 0, \dots, h_x = 0$ whose determinant $\Delta = (ab \dots h) \neq 0$. Multiplying each equation by the co-factor in Δ of the coefficient of x_i, and adding, we get $x_i \Delta = 0$. Hence x_i vanishes for each value of i, and only the zero solution $x_1 = x_2 = \dots = x_n = 0$ exists.

Next if $r = n - 1$, so that there must be $n - 1$ independent rows, there is exactly one solution for the ratios $x_1 : x_2 : \dots : x_n$, namely

$$x_1 : x_2 : \dots : x_n = K_1 : K_2 : \dots : K_n,$$

where these K's are the $n - 1$-rowed determinants in the $n - 1$ by n matrix of these $n - 1$ equations. This follows in the same way by multiplying the $n - 1$ equations respectively by the co-factors of $a, b \dots$ in the determinant $(ab \dots g)_{123\dots n-1}$, and adding.

Further if $r < n-1$, we can solve the r equations for just $x_1, x_2, \dots, x_r$ in terms of the remaining $n-r$ homogeneous variables $x_{r+1}, \dots, x_n$, again by use of co-factors of elements in a column of the non-vanishing determinant

$$(ab \dots g)_{12\dots r}.$$

EXAMPLES

1. The rank of the system of equations

$$\begin{aligned} x + y + z + t &= 0 \\ 2x + 2y + 3z - t &= 0 \\ z - 3t &= 0 \\ 3x + 3y + 5z - 3t &= 0 \end{aligned}$$

is two. We could express two of x, y, z, t in terms of the others, but not *any* two. We must exclude the case, x, y.

2. Given two equations

$$\begin{aligned} a_1x_1 + b_1x_2 + c_1x_3 + d_1x_4 &= 0 \\ a_2x_1 + b_2x_2 + c_2x_3 + d_2x_4 &= 0 \end{aligned}$$

the most general linear equation derivable is

$$(ap)x_1 + (bp)x_2 + (cp)x_3 + (dp)x_4 = 0$$

where $(ap) = a_1p_2 - a_2p_1$ and p_1, p_2 are arbitrary. We eliminate x_1 or x_2 or x_3 or x_4 by putting $p = a, b, c, d$ in succession, provided not all of the determinants (ab) vanish.

3. Given three equations

$$\begin{aligned} a_1x_1 + b_1x_2 + c_1x_3 + d_1x_4 &= 0 \\ a_2x_1 + b_2x_2 + c_2x_3 + d_2x_4 &= 0 \\ a_3x_1 + b_3x_2 + c_3x_3 + d_3x_4 &= 0 \end{aligned}$$

prove $(apq)x_1 + (bpq)x_2 + (cpq)x_3 + (dpq)x_4 = 0$ where $(apq) = \Sigma \pm a_1p_2q_3$, and the six elements p_i, q_j are arbitrary. By putting p, q equal to two of a, b, c, d we eliminate two of the x's from the equations, provided not all determinants (abc) vanish.

4. Prove similarly that $r-1$ unknowns x_i can be eliminated from r such equations whose matrix is of rank r.

5. Prove the fundamental identity

$$(abcq)p_x - (abcp)q_x = (bcpq)a_x + (capq)b_x + (abpq)c_x,$$

where $p_x = p_1x_1 + p_2x_2 + p_3x_3 + p_4x_4$. If the matrix $[abcd]$ has rank two, prove $(abcp) = 0$, $(abcq) = 0$.

This fundamental identity now explicitly indicates the linear relation between three equations $a_x = 0$, $b_x = c_x = 0$, whose rank is two.

3. The Upper Suffix Notation.

Let us denote the elements of the reciprocal determinant Δ^{-1} by a^1, a^2, a^3, . . . , m^n, in defiance of previous practice which reserves this notation for powers of a. The context will make it clear what is meant. So at present we shall understand the indices 1, 2, 3, . . . , n to be distinguishing marks like the accents which are probably more familiar

$$a,\ a',\ a'',\ \ldots.$$

By this means we can exhibit a wonderful parallelism running through the whole theory of determinants, starting with the obvious pair of conditions

$$\begin{aligned} \Delta\ \ &= \Sigma \pm a_1 b_2 \ldots m_n \\ \Delta^{-1} &= \Sigma \pm a^1 b^2 \ldots m^n. \end{aligned}$$

Since $A_i \div \Delta$ is by definition a^i, and since

$$\Delta = a_1 A_1 + \ldots + a_n A_n,$$

we have

$$1 = a_1 a^1 + a_2 a^2 + \ldots + a_n a^n.$$

If A^i is the co-factor of a^i in Δ^{-1} we find $A^i \times \Delta = a_i$. This is soon proved. For, multiplying rows by rows, $A^1 \times \Delta$

$$\begin{vmatrix} 1 & . & . & . & \ldots \\ a^2 & b^2 & c^2 & d^2 & \ldots \\ a^3 & b^3 & c^3 & d^3 & \ldots \\ a^4 & b^4 & c^4 & d^4 & \ldots \\ . & . & . & . & . \ . \ . \end{vmatrix} \begin{vmatrix} a_1 & b_1 & c_1 & d_1 & .. \\ a_2 & b_2 & c_2 & d_2 & .. \\ a_3 & b_3 & c_3 & d_3 & .. \\ a_4 & b_4 & c_4 & d_4 & .. \\ . & . & . & . & . \ . \ . \end{vmatrix}$$

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 & \ldots \\ . & 1 & . & . & \ldots \\ . & . & 1 & . & \ldots \\ . & . & . & 1 & \ldots \\ . & . & . & . & . \ . \ . \end{vmatrix} = a_1,$$

and similarly for a_i.

This leads to a general theorem, due to Jacobi:[1]

Each minor of Δ is proportional to the corresponding complementary co-factor of Δ^{-1}, the ratio being Δ.

The full significance of this remarkable theorem is best seen by taking a particular case, when $\Delta = | a_1 b_2 c_3 d_4 |$, and $\Delta^{-1} = | a^1 b^2 c^3 d^4 |$,

---

[1] Cf. *Crelle*, **12** (1834), 9.

$$\frac{a_1}{|b^2 c^3 d^4|} = \frac{b_1}{|c^2 a^3 d^4|} = \frac{c_1}{|a^2 b^3 d^4|} = \frac{d_1}{|a^2 c^3 b^4|} = \frac{|a_1 b_2|}{|c^3 d^4|} = \frac{|b_1 c_2|}{|a^3 d^4|} = \&\text{c}.$$

$$= \frac{|a_1 b_2 c_3|}{d^4} = \frac{|a_1 b_2 c_3 d_4|}{1} = \frac{1}{|a^1 b^2 c^3 d^4|}.$$

There are altogether 70 different equal ratios involved here, while each ratio involves all the letters and all the digits 1, 2, 3, 4. In general we have

$$\frac{|a_i b_j \ldots f_k|}{|g^l \ldots m^n|} = \frac{|a_1 b_2 \ldots m_n|}{1} = \frac{1}{|a^1 b^2 \ldots m^n|},$$

where, to fix the sign, both letter and suffix rows are algebraic complements

$$\begin{matrix} ab \ldots f, & gh \ldots m, \\ ij \ldots k, & l \ldots n, \end{matrix}$$

and the partition is taken in every possible way.

The proof is immediate for any particular case, by the multiplication theorem. Thus $| a_1 b_2 | : | c^3 d^4 | = | a_1 b_2 c_3 d_4 | : 1$ since

$$|c^3 d^4| \quad |a_1 b_2 c_3 d_4| = \begin{matrix} \cdot & 1 & \cdot & \cdot \\ a^3 & b^3 & c^3 & d^3 \\ a^4 & b^4 & c^4 & d^4 \end{matrix} \qquad \begin{matrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{matrix}$$

$$\begin{matrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ \cdot & \cdot & 1 & \cdot \\ \cdot & \cdot & \cdot & 1 \end{matrix} \qquad \begin{matrix} a_1 & a_2 \\ b_1 & b_2 \end{matrix} = | a_1 b_2 |.$$

The only difficulty in this proof is to decide how to make the minor look like an n-rowed determinant. An easy way to remember what to do is to notice how two complementary blocks of the unit determinant,

$$|\delta_{ij}| = \left| \begin{matrix} 1 & \cdot & \ldots \\ \cdot & 1 & \ldots \\ \cdot & \cdot & \cdot \quad \cdot \\ \cdot & \cdot & \ldots \end{matrix} \right.$$

are utilized in the course of the work.

To prove $| c^1 a^4 | \, | a_1 b_2 c_3 d_4 | = | b_2 d_3 |$ where the order c, a, b, d of the letters has been disturbed, write $| a_1 b_2 c_3 d_4 |$ as $| c_1 a_2 b_3 d_4 |$ and proceed as before.

It is important to recognize this theorem in the form of the *adjugate determinant and its minors.* Thus if capital letters denote co-factors of small letters in Δ, the following would be typical when $\Delta = | a_1 b_2 c_3 d_4 |$:

$$\begin{aligned} | a_1 b_2 | \, \Delta &= | C_3 D_4 | \\ c_2 \Delta^2 &= | A_3 B_1 D_4 | \\ \Delta^3 &= | A_1 B_2 C_3 D_4 | , \end{aligned}$$

and so on.

4. **The Theorem of Corresponding Matrices** (the Binet Cauchy Theorem).

Next consider two arbitrary square matrices A and X of order n,

$$A = \begin{bmatrix} a_1 & \dots & m_1 \\ \cdot & \cdot \quad \cdot & \cdot \\ a_n & \dots & m_n \end{bmatrix} \qquad X = \begin{bmatrix} x_1 & \dots & t_1 \\ \cdot & \cdot \quad \cdot & \cdot \\ x_n & \dots & t_n \end{bmatrix},$$

together with their transposed matrices A' and X'. It is assumed that the determinant $| A | \neq 0$.

By the same device of interchanging corresponding rectangular subsections of these we obtain another important theorem, which is illustrated well enough again by taking $n = 4$.

First replace the top row of A' by that of X', raise the suffixes of other rows and multiply the determinant so found by $| A |$. Then

$$\begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ b^1 & b^2 & b^3 & b^4 \\ c^1 & c^2 & c^3 & c^4 \\ d^1 & d^2 & d^3 & d^4 \end{vmatrix} \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = \begin{vmatrix} a_x & b_x & c_x & d_x \\ & & & \\ & & & \\ & & & \end{vmatrix} = a_x \qquad (8)$$

which is obvious otherwise by expanding the first determinant by its top row and using the previous theorem. In fact

$$\begin{aligned} &(x_1 | b^2 c^3 d^4 | + x_2 | b^3 c^1 d^4 | + \&c.) \, | a_1 b_2 c_3 d_4 | \\ &= x_1 a_1 + x_2 a_2 + x_3 a_3 + x_4 a_4 \\ &= a_x \end{aligned}$$

Next do the same for the two top rows of A' and X'. This gives

$$\begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ c^1 & c^2 & c^3 & c^4 \\ d^1 & d^2 & d^3 & d^4 \end{vmatrix} \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} \begin{matrix} a_x & b_x & c_x & d_x \\ a_y & b_y & c_y & d_y \\ & & 1 & \\ & & & 1 \end{matrix} = a_x b_y - a_y b_x. \quad (9)$$

This last is so important a form that it has a special notation

$$a_x b_y - a_y b_x = (ab \mid xy). \quad . \quad . \quad . \quad . \quad (10)$$

Again expand the first determinant in (9), this time by its two top rows, and use the previous theorem. It gives

$$\begin{aligned} &(\Sigma\,(xy)_{12}\,(cd)^{34})\,(abcd) \\ &= \Sigma\,(xy)_{12}\,(ab)_{12} \end{aligned}$$

summed to six terms ($= 4\,!/2\,!\,2\,!$). Thus

$$a_x b_y - a_y b_x = (ab \mid xy) = \Sigma\,(xy)_{ij}\,(ab)_{ij}, \quad . \quad (11)$$

and as this proof equally well applies when Δ is of order n, and there are $n - 2$ rows below the top two, we may sum this last for

$$\left.\begin{matrix} i = 1, 2, 3, \ldots, n, \\ j = 1, 2, 3, \ldots, n, \end{matrix}\right\} \quad i \neq j.$$

Similarly for three top rows

$$\begin{vmatrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \\ d^1 & d^2 & d^3 & d^4 \end{vmatrix} \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & \\ a_4 & b_4 & c_4 & \end{vmatrix} = \begin{vmatrix} a_x & b_x & c_x & d_x \\ a_y & b_y & c_y & d_y \\ a_z & b_z & c_z & d_z \\ . & . & . & 1 \end{vmatrix} \quad (12)$$

$$= \Sigma \pm a_x b_y c_z = (abc \mid xyz), \text{ say.}$$

Also on expanding the first determinant by its three top rows, it gives

$$\begin{aligned} &(\,(xyz)_{123}\,d^4 + \&\text{c.})\,(abcd) \\ &= (xyz)_{123}\,(abc)_{123} + \&\text{c.} \\ &= \Sigma\,(xyz)_{ijk}\,(abc)_{ijk} \end{aligned}$$

to four terms ($= 4\,!/3\,!\,1\,!$). Thus

$$\left.\begin{aligned} \Sigma \pm a_x b_y c_z = (abc \mid xyz) = \Sigma\,(abc)_{ijk}\,(xyz)_{ijk}\,. \\ i, j, k = 1, 2, 3, \ldots, n. \\ i \neq j, \;\; i \neq k, \;\; j \neq k. \end{aligned}\right\} \quad (13)$$

Lastly, if we remove all rows of A' we revert to the ordinary product theorem of determinants:

$$\left.\begin{matrix} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \\ z_1 & z_2 & z_3 & z_4 \\ t_1 & t_2 & t_3 & t_4 \end{matrix}\right| \begin{vmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & & c_2 & d_2 \\ a_3 & & c_3 & d_3 \\ a_4 & b_4 & c_4 & d_4 \end{vmatrix} = \begin{vmatrix} a_x & b_x & c_x & d_x \\ a_y & b_y & c_y & d_y \\ a_z & b_z & c_z & d_z \\ a_t & b_t & c_t & d_t \end{vmatrix} \qquad (14)$$

$= (abcd \mid xyzt)$ say.

We collect these results, which will turn out to be very significant, and now have the following n identities involving elements of two square matrices of order n:

$$a_x = (a \mid x) = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n, \quad . \quad . \quad . \quad . \quad (15_1)$$

$$\begin{vmatrix} a_x & b_x \\ a_y & b_y \end{vmatrix} = (ab \mid xy) = \Sigma\, (ab)_{ij}\, (xy)_{ij}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15_2)$$

$$\begin{vmatrix} a_x & b_x & c_x \\ a_y & b_y & c_y \\ a_z & b_z & c_z \end{vmatrix} = (abc \mid xyz) = \Sigma\, (abc)_{ijk}\, (xyz)_{ijk}, \quad . \quad . \quad (15_3)$$

$$\begin{matrix} a_x\, b_x\, c_x \ldots m_x \\ a_y\, b_y\, c_y \ldots m_y \\ \\ a_t\, b_t\, c_t \ldots m_t \end{matrix} = (abc\ldots m \mid xyz\ldots t) = (abc\ldots m)\,(xyz\ldots t). \quad (15_n)$$

$$i, j, \ldots = 1, 2, 3, \ldots, n.$$

Here there are $\binom{n}{r}$ terms on the right of the rth identity, obtained by choosing the r suffixes in all different combinations from among 1, 2, 3, . . . , n.

Determinants of this type on the left, but of higher order than n, vanish identically. For by §12, (V), p. 50, the elements of the $(n + 1)$th column are each the same linear function of the first n columns. Hence the $(n + 1)$th column is linearly related to these columns.

At the outset we have assumed that $| A | \neq 0$, but the above results hold for all values of the elements $a_i, b_j, \ldots, x_i, \ldots$ concerned. For if in the rth identity the r columns $a, b, \ldots, k$ are linearly related, then

$$\theta a_i + \phi b_i + \ldots + \omega k_i = 0, \qquad i = 1, 2, \ldots, n.$$

Multiplying by x_i and summing for i, we have

$$\theta\, a_x + \phi\, b_x + \dots + \omega\, k_x = 0,$$

so that $a_x,\ b_x,\ \dots,\ k_x$ are also linearly related. Consequently both sides of relation 15_r vanish identically.

But if $a, b, \dots, k$ are not linearly related, at least one r-rowed determinant of the matrix $[a, b, \dots, k]$ does not vanish, $|A_r|$ say. Let $[l \dots m]$ denote the $(n - r)$-columned matrix, formal dual to $[a, b, \dots, k]$ in $A = [ab \dots m]$. By choosing $[l \dots m]$ to be zero except for its diagonal elements complementary to $|A_r|$ which we take as units, we make

$$|A| = |A_r| \neq 0.$$

Thus the original assumption is covered.

These results may be put into a slightly different form, of great importance in the theory of matrices.

If A *is any matrix of* n *columns, and* B *is any matrix of* n *rows, any* r*-rowed determinant* D *of the product matrix* AB *is equal to a sum of terms each a product of an* r*-rowed determinant of* A *and an* r*-rowed determinant of* B.

For this is a statement of formula (15_r) when

$$A = \begin{matrix} a_1 & a_2 & & a_n \\ b_1 & b_2 & & b_n \end{matrix} \qquad B = \begin{matrix} x_1 & y_1 & \dots \\ x_2 & y_2 & \dots \\ \cdot & \cdot & \cdot \\ x_n & y_n & \dots \end{matrix}$$

Here A and B need not be square matrices.

5. Inner Product of Two Rectangular Matrices. Bideterminant.

The determinants $(a|x)$, $(ab|xy)$, $(abc|xyz)$, ... elaborated in the last article have many useful properties. It will be seen by interchanging any pair among $abc\dots$ or among $xyz\dots$ that the sign of the whole is changed; for this amounts to an interchange of rows in the second or first factor of the original product. Thus

$$(abc\,|\,xyz) = -(bac\,|\,xyz) = -(abc\,|\,yxz) = \dots .$$

Such a property is summed up by saying that $(abc\,|\,xyz)$ *alternates in both* abc *and* xyz. Further, it is symmetrical in these two groups; thus

$$(abc\,|\,xyz) = (xyz\,|\,abc).$$

Once more, it has exactly the same suffixes ijk for abc and for xyz in any particular term of its expansion

$$\Sigma\, (abc)_{ijk}\, (xyz)_{ijk}.$$

For this reason it is an inner product of two sets of $\binom{n}{3}$ quantities, namely the set of determinants of the three-line matrix

$$A_3' = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 & \dots & a_n \\ b_1 & b_2 & b_3 & b_4 & \dots & b_n \\ c_1 & c_2 & c_3 & c_4 & \dots & c_n \end{bmatrix}$$

and the corresponding set

$$\Theta_3' = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 & \dots & x_n \\ y_1 & y_2 & y_3 & y_4 & \dots & y_n \\ z_1 & z_2 & z_3 & z_4 & \dots & z_n \end{bmatrix}$$

Likewise for each of these functions $(a \mid x)$, $(ab \mid xy)$, ... The typical one can be called rth compound inner product, or the inner product of two r by n matrices, r giving the number of different symbols before or after the vertical line. Further, these matrices are subdivisions by rows of the transposed of A', X' the square arrays from which we started. Let such inner products be called *bideterminants.*

6. Laplace Developments of the Inner Products.

First we may write any such bideterminant as

$$(abc \mid xyz) = \Sigma \pm a_x b_y c_z,$$

and in the notation of p. 27 this can be written

$$(abc \mid xyz) = \dot{a}_x \dot{b}_y \dot{c}_z = a_{\dot{x}} b_{\dot{y}} c_{\dot{z}},$$

where either a, b, c are deranged, or xyz. By a $(2 \mid 1)$ Laplace development this is also

$$(ab \mid yz)\, c_x + (ab \mid zx)\, c_y + (ab \mid xy)\, c_z,$$

and this principle may be extended to any order involving r pairs of letters a, x; b, y; &c., if $r \leqslant n$.

7. Rank of the Product of Matrices.

The rank of the product of two matrices cannot exceed the rank of either factor.

For in the result of §4, p. 81, if all r-rowed determinants of A (or of B) are zero, the same is true of all r-rowed determinants of their product AB. Again

The rank of a matrix A *of* m *rows and* n *columns is unaltered by multiplying* A *fore or aft by a conformable non-singular square matrix.*

For if r is the rank of A, and B is a square non-singular matrix of order n, the rank ρ of AB has just been proved to be not greater than r. Likewise the rank of $(AB)B^{-1}$ cannot exceed ρ, that of its first factor. Hence $\rho = r$. Similarly for a product with A as after factor.

8. The Simplex.

We now consider a theorem analogous to Jacobi's theorem, which may first be illustrated by the case of four rows and columns. Suppose we have, as before, two double sets, say

$$\begin{array}{cccc} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \end{array} \qquad \begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \\ y_1 & y_2 & y_3 & y_4 \end{array}$$

so chosen that the matrix

$$\begin{bmatrix} a_x & a_y \\ b_x & b_y \end{bmatrix}$$

vanishes identically: thus $a_x = a_y = b_x = b_y = 0$, or, in full,

$$a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 = 0,$$
$$b_1 x_1 + b_2 x_2 + b_3 x_3 + b_4 x_4 = 0,$$

and similarly for y. Eliminating x_1 we obtain

$$(ab)_{12}\, x_2 + (ab)_{13}\, x_3 + (ab)_{14}\, x_4 = 0.$$

Similarly

$$(ab)_{12}\, y_2 + (ab)_{13}\, y_3 + (ab)_{14}\, y_4 = 0.$$

These are two equations for the three quantities $(ab)_{12}$, $(ab)_{13}$, $ab)_{14}$. Hence, by solving,

$$(ab)_{12} : (ab)_{13} : (ab)_{14} = (xy)_{34} : (xy)_{42} : (xy)_{23},$$

and for similar reasons, by eliminating x_2 or x_3 or x_4 originally, we obtain

$$\frac{(ab)_{12}}{(xy)_{34}} = \frac{(ab)_{13}}{(xy)_{42}}. \quad \frac{(ab)_{14}}{(xy)_{23}} = \frac{(ab)_{23}}{(xy)_{14}} = \frac{(ab)_{42}}{(xy)_{13}} = \frac{(ab)_{34}}{(xy)_{12}}.$$

These formulæ, which are fundamental in the study of line geometry in threefold space, are typical of a general set involving complementary rectangles P and π, say, from two square matrices of order n. As an illustration let $n = 5$, and let the matrix

$$\begin{bmatrix} a_x & a_y & a_z \\ b_x & b_y & b_z \end{bmatrix}$$

vanish identically. Here five sets a, b, x, y, z are involved, and

$$a_x \equiv a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4 + a_5 x_5 = 0.$$

As before we obtain, by eliminating x_1, y_1, z_1 from the three pairs of equations,

$$\begin{aligned} (ab)_{12}\, x_2 + (ab)_{13}\, x_3 + (ab)_{14}\, x_4 + (ab)_{15}\, x_5 &= 0, \\ (ab)_{12}\, y_2 + (ab)_{13}\, y_3 + (ab)_{14}\, y_4 + (ab)_{15}\, y_5 &= 0, \\ (ab)_{12}\, z_2 + (ab)_{13}\, z_3 + (ab)_{14}\, z_4 + (ab)_{15}\, z_5 &= 0, \end{aligned}$$

which are just enough equations to determine the ratios $(ab)_{12} : (ab)_{13}$, &c., in terms of x, y, z. Multiplying these equations respectively by $(yz)_{45}$, $(zx)_{45}$, $(xy)_{45}$, and adding, we deduce

$$(ab)_{12}\, (xyz)_{245} + (ab)_{13}\, (xyz)_{345} = 0,$$

or

$$(ab)_{12} : (ab)_{13} = (xyz)_{345} : (xyz)_{425};$$

and by symmetry

$$\frac{(ab)_{12}}{(xyz)_{345}} = \frac{(ab)_{13}}{(xyz)_{425}} = \ldots = \frac{(ab)_{ij}}{(xyz)_{klm}},$$

where ij, klm are algebraic complements of 12345.

In general let there be r sets $a, b, \ldots, k$ and $n - r$ sets $x, y, \ldots, t$ such that the inner product of any two sets chosen from these different groups vanishes. Then the determinants of these two matrices are proportional, namely

$$\frac{(ab \ldots k)_{\alpha\beta \ldots \delta}}{(xy \ldots t)_{\lambda\mu \ldots \rho\sigma}} = \frac{(ab \ldots k)_{\alpha'\beta' \ldots \delta}}{(xy \ldots t)_{\lambda\mu' \ldots \rho'\sigma}}$$

where $\alpha\beta \dots \delta$, $\lambda\mu \dots \rho o$ are algebraic complements of $123 \dots n$, and so are the accented suffixes.

Lastly consider the double set arranged in two rows each of n letters,

$$\begin{array}{cccccc} a & b & c & \dots & l & m \\ x & y & z & \dots & s & t \end{array}$$

where each letter denotes a *vector*, say

$$a = [a_1, a_2, \dots, a_n],$$
$$x = \{x_1, x_2, \dots, x_n\}$$

and the inner product of any two vectors not in the same row or column vanishes. When none of the inner products a_x, b_y, $c_z, \dots, m_t$ vanish, such a system of vectors is called a *simplex* of the nth category: the vectors of one row, it matters not which, are called *points*, and those of the other row *primes*. Further, let $p_2, p_3, \dots, p_{n-1}$ denote the sets of determinants $(ab)_{ij}$, $(abc)_{ijk} \dots$ respectively, and $\pi_2, \pi_3, \dots, \pi_{n-1}$ the sets $(xy)_{ij}$, $(xyz)_{ijk}, \dots$ Then the preceding results can be written in the form

$$\begin{array}{llcll} x = \pi_1 & x_\alpha & = & k(bc \dots m)_{\beta\gamma\dots} & p_{n-1} \\ \pi_2 & (xy)_{\alpha\beta} & = & k'(c \dots m)_{\gamma \dots \sigma} & p_{n-2} \\ \pi_3 & (xyz)_{\alpha\beta\gamma} & = & k''(d \dots m)_{\delta \dots \epsilon} & p_{n-3} \\ & \cdot\;\cdot\;\cdot & \cdot & \cdot\;\cdot\;\cdot\;\cdot\;\cdot & \cdot\;\cdot \\ \pi_{n-1} & (xyz \dots s)_{\alpha\beta\gamma \dots \rho} & = & k^{(n-1)} m_\sigma & p_1 = m \end{array}$$

where the suffixes are algebraic complements of $123 \dots n$, and the coefficients k are constant for all such suffixes.

EXAMPLES

1. If $\Delta = (abc \dots m)$ and $D = (xyz \dots t)$ denote the determinants of the above double set of vectors, prove $D\Delta = a_x b_y c_z \dots m_t$.

2. Prove that each column of Δ is in fact a multiple of the corresponding column of co-factors of D, and vice versa.

3. Exhibit the above relations when $n = 3$, and interpret them when x, y, z are three vertices of a triangle whose sides are a, b, c.

4. If $n = 4$, show that the simplex is a tetrahedron of four vertices x, y, z, t, four planes a, b, c, d, and six lines.

Here $b_\xi = 0$ is the homogeneous equation of a plane in point co-ordinates $\{\xi_1, \xi_2, \xi_3, \xi_4\}$ or of a point in plane co-ordinates $[b_1, b_2, b_3, b_4]$. The conditions imply that points y, z, t but *not* x are on plane a; and so on cyclically. The set

$$p = [(ab)_{12}, (ab)_{13}, (ab)_{14}, (ab)_{23}, (ab)_{42}, (ab)_{34}]$$

is proportional to the set

$$\pi = [(xy)_{34}, (xy)_{42}, (xy)_{23}, (xy)_{14}, (xy)_{13}, (xy)_{12}]$$

as in §6 above. The set p is called the set of *axial* co-ordinates of the line common to planes a, b: the set π forms the set of *line* co-ordinates. These p and π sets are often taken to be identically equal since the introduction of a non-zero common factor to homogeneous co-ordinates does not alter the actual point, line or plane represented.

9. Extended Form of Cauchy's Theorem, commonly called Sylvester's Theorem on Compound Determinants.

The theorem that $| a_1 b_2 \ldots m_n |^{n-1} = | a^1 b^2 \ldots m^n |^{1-n} = | A_1 B_2 \ldots M_n |$ is a special case of a remarkable general theorem virtually due to Cauchy.[1] Let $|(ab)_{ij}|$ denote the determinant of order $\binom{n}{2}$, with the ${}_nC_2$ combinations like ab to typify columns and the ${}_nC_2$ like ij to typify rows. Let us call this the *second compound* of Δ, and denote it by D_2. Thus if $n = 3$

$$D_2 = |(ab)_{ij}| = \begin{vmatrix} (ab)_{12} & (ac)_{12} & (bc)_{12} \\ (ab)_{13} & (ac)_{13} & (bc)_{13} \\ (ab)_{23} & (ac)_{23} & (bc)_{2} \end{vmatrix}$$

To avoid ambiguity let the alphabetical and the ascending order of letters and suffixes be maintained. If $n = 4$ there are six rows and columns.

Similarly let $D_3 = |(abc)_{ijk}|$ be the third compound of Δ, denoting the corresponding determinant of order $\binom{n}{3}$. Thus if $n = 4$ its leading diagonal is

$$(abc)_{123} \quad (abd)_{124} \quad (acd)_{134} \quad (bcd)_{234}.$$

And so on until finally Δ itself is reached. The Cauchy-Sylvester theorem is this:

Each determinant $\mathrm{D}_r = \left|(ab\ldots)_{ij\ldots}\right|$ *is a positive integral power of* Δ, *the power for the* r*th compound being* $\binom{n-1}{r-1}$.

This is proved by considering an adjoint determinant whose elements are co-factors, in the original determinant Δ, of the elements $(ab\ldots)_{ij\ldots}$.

---

[1] See Muir, *History of Determinants*, **1**, 118.

Let $(ab\ldots f)_{ij\ldots k}$, $(gh\ldots m)_{l\ldots n}$ be such co-factors. We form the two determinants

$$|(ab\ldots)_{ij\ldots}|, \quad |(gh\ldots)_{l\ldots}|,$$

and multiply them together, column by column, not row by column. The resulting inner products which go to form elements of the product determinant are all determinants of order n, because the inner products are actual Laplace developments of these determinants. The leading diagonal determinants are all equal to Δ, and the others vanish as in the original Cauchy theorem.

To illustrate this let $n=5$, $\Delta=(abcde)$ and the co-factor determinants be $|(ab)_{ij}|$, $|(cde)_{klm}|$. Then

$$\begin{array}{ccc} (ab)_{12} & (ac)_{12} & \cdots \\ (ab)_{13} & (ac)_{13} & \cdots \\ \cdot\;\cdot\;\cdot & & \end{array} \times \begin{array}{cc} (cde)_{345} & (bde)_{345} \\ (cde)_{245} & (bde)_{245} \end{array}$$

$$\begin{array}{cc} (abcde) & (abbde) \\ (accde) & (abcde) \end{array} \qquad\qquad \Delta \qquad\qquad = \Delta^{10}.$$

In general the result of multiplying these two determinants each of order $\binom{n}{r}$ is $\Delta^{\binom{n}{r}}$.

But Δ has no polynomial factors (§4, p. 33); hence $\Delta^{\binom{n}{r}}$ also has none. Accordingly both determinants on the left are powers of Δ, to a numerical factor. Taking the special case when Δ is $|\delta_{ik}|$, the unit determinant, the numerical factor is seen to be unity: and further, since, in the left-hand factor determinant, the letter a enters into the columns, and therefore into each term in its expansion $\binom{n-1}{r-1}$ times, it follows that

$$D_r = \Delta^{\binom{n-1}{r-1}}.$$

The other factor must be D_{n-r}, which is equal to $\Delta^{\binom{n-1}{n-r-1}}$. For by the theory of indices,

$$D_r D_{n-r} = \Delta^{\binom{n-1}{r-1}} \Delta^{\binom{n-1}{n-r-1}} = \Delta^{\binom{n}{r}}.$$

We may now summarize this Cauchy-Sylvester theorem for all compound determinants D_i, derived from a given determinant of order n as follows:

$$\begin{aligned} D_1 &= \Delta \\ D_2 &= \Delta^{\binom{n-1}{1}} \\ D_3 &= \Delta^{\binom{n-1}{2}} \\ & \\ D_{n-1} &= \Delta^{\binom{n-1}{n-2}} = \Delta^{n-1} \\ D_n &= \Delta. \end{aligned}$$

Thus for $n = 5$ this set gives Δ, Δ^4, Δ^6, Δ^4, Δ.

EXAMPLES

1. Prove the analogous reciprocal results for upper suffixes:

$$|(ab)^{ij}| = 1/D_2 = \Delta^{1-n}, \quad |(abc)^{ijk}| = 1/D_3, \quad \&c.$$

2. Extend the Jacobi ratio theorem to cover the case of minors of reciprocal adjugate determinants, e.g. $(ab)_{ij}|$ and $|(ab)^{ij}|$.

10. The Generalized Ratio Theorem.

Starting with the Jacobi ratio theorem (§3, p. 78), which the quaternary case illustrates,

$$|a_1 b_2 c_3 d_4| = \frac{a_1}{|b^2 c^3 d^4|} = \frac{|a_1 b_2|}{|c^3 d^4|} = \ldots = \frac{1}{|a^1 b^2 c^3 d^4|},$$

we can adjoin further equal ratios as follows. Take any number of arbitrary sets

$$\begin{matrix} \{x^1 & x^2 & \ldots & x^n\}, \\ \{y^1 & y^2 & \ldots & y^n\}, \end{matrix}$$

and multiply single-letter numerators of suffix i together with their denominators, by x^i; double letter numerators $|a_i b_j|$ together with their denominators by $|x^i y^j|$; and so on. Then sum the numerators forming $\Sigma a_i x^i$, and sum their corresponding denominators $\Sigma x^i |b^j c^k d^l|$. Also sum $\Sigma |ab||xy|$ and their denominators. The result is to obtain further ratios equal to the original, such as

$$\frac{(x|a)}{|x^1 b^2 c^3 d^4|}.$$

And if by $(xcbd)'$ we understand this *upper* suffix determinant in contrast to $(xbcd)$ which means $|\,x_1\,b_2\,c_3\,d_4\,|$, we have

$$|\,a_1 b_2 c_3 d_4\,| = (abcd) = \frac{(x\,|\,a)}{(xbcd)'} = \frac{(x\,|\,b)}{(xcad)'} = \frac{(xy\,|\,ab)}{(xycd)'} = \dots = \frac{1}{(abcd)}$$

for all values of x, y, ...

A correlative statement, involving arbitrary sets $[u_1, u_2, \dots, u_n]$, $[v_1, v_2, \dots, v_n]$, . . . , with lower suffixes, follows in exactly the same way:

$$(abcd) = \frac{(uvwa)}{(uvw\,|\,bcd)'} = \frac{(uvwb)}{(uvw\,|\,cad)'} = \frac{(uvab)}{(uv\,|\,cd)'} = \dots = \frac{1}{(abcd)'}.$$

11. **Tensor Constants of the Fundamental Identities** (pp. 44–48).

In the fundamental identities certain matrices R, L, M, N enter. The columns of some are deranged from term to term, those of others maintain their relative position. We shall call these latter matrices the *constants* of the identity concerned, while the others, resolved into their ultimate columns, will be called the *variable vectors* or simply the *variables*.

Thus, for example, in §8 (30), p. 43, which leads to identity (31) namely to

$$(abc\dot{d})\,(\dot{e}\dot{f}gh) = (\dot{a}def)\,(\dot{b}\dot{c}gh), \tag{16}$$

the constant is

$$[gh] = \begin{bmatrix} g_1 & h_1 \\ g_2 & h_2 \\ g_3 & h_3 \\ g_4 & h_4 \end{bmatrix} \tag{17}$$

Any identity arising from deranging a product of p such determinants, as in §11, (I), (II), (III), p. 48, evidently has $p - 1$ constants M, N, ... They have respective currencies j, k, ...

Now consider the above example, with each second factor developed by a Laplace expansion as

$$(efgh) = \Sigma\,(ef)_{pq}\,(gh)_{rs}. \quad . \quad . \quad . \quad . \tag{18}$$

$$p, q, r, s = 1, 2, 3, 4.$$

Since (16) is an identity for all values of g_r, h_r, in particular, we may put $g_r = h_s = 1$ with all other six elements g and h zero. Hence

$$(abc\dot{d})\,(\dot{e}\dot{f})_{pq} = (def\dot{a})\,(\dot{b}\dot{c})_{pq} \quad . \; . \; . \quad (19)$$

where p, q are any two among 1, 2, 3, 4.

Now let a set of six arbitrary quantities be chosen

$$G = [G_{12}, G_{13}, G_{14}, G_{34}, G_{42}, G_{23}] = [G_{rs}]. \quad . \quad (20)$$

Multiply (19) through by G_{rs} and sum for the six sets of values of pq, rs as given by (18), finally writing

$$(ef\,G) \text{ for } \Sigma\,(ef)_{pq}\,G_{rs}.$$

Then

$$(abc\dot{d})\,(\dot{e}\dot{f}G) = (def\dot{a})\,(\dot{b}\dot{c}G). \qquad (21)$$

Here we have a slightly different form of the identity. It now involves a constant set G defined by six elements as in (20), but not necessarily by eight original elements as in (17).

In exactly the same way the identities (I), (II), (III), §11, p. 48, may be treated. The constant matrix M of currency j may be replaced by an arbitrary set of $\binom{n}{j}$ elements M

$$M = [M_{12\ldots j},\ \ldots] = [M_{r_1 r_2 \ldots r_j}]. \qquad (22)$$

Similarly for the other possible constants in the identity.

And again, returning to identity (19), if we multiply each member by an arbitrary quantity H^{pq}, taking

$$H = [H^{34}, H^{42}, H^{23}, H^{12}, H^{13}, H^{14}] = [H^{pq}], \quad . \quad (23)$$

and if $(ef \mid H)$ means $\Sigma\,(ef)_{pq}\,H^{pq}$, we may write the identity as

$$(abc\dot{d})\,(\dot{e}\dot{f} \mid H) = (def\dot{a})\,(\dot{b}\dot{c} \mid H). \quad . \; . \; . \quad (24)$$

Thus (16), (19), (21), (24) are essentially the same identity as far as the *variables* a, b, c, d, e, f are concerned.

Definition of Tensor.—*For a given category* n, *a set of quantities* $[\mathrm{M}_{\mathrm{r_1 r_2 \ldots r_j}}]$ *where the suffixes take all values* 1, 2, . . . , n *is called a tensor of order* j. A vector is a tensor of order unity.

12. Application of the Principle of Duality.

Let us make formal duals (§13, p. 54) of the vectors $a, b, c, \ldots$: in other words we consider each vector to furnish a column of a perfectly arbitrary n-rowed determinant, and take as formal dual of a the set of n co-factors of the column $\{a_1, a_2, \ldots, a_n\}$. Then if this determinant is written $\Sigma a_i \alpha^i \equiv (a \mid \alpha)$, the set

$$[\alpha^1, \alpha^2, \ldots, \alpha^n]$$

is formal dual of

$$\{a_1, a_2, \ldots, a_n\}.$$

Now let Greek letters α, β, γ be formal duals of a, b, c. Then there will be a compound n-rowed determinant

$$(\alpha\beta\gamma \ldots \mu) = \Sigma \pm \alpha^1 \beta^2 \gamma^3 \ldots \mu^n,$$

formally dual to

$$(abc \ldots m) = \Sigma \pm a_1 b_2 c_3 \ldots m_n.$$

Manifestly fundamental identities exist among such determinants. Thus if $n = 4$ we have as dual of (16)

$$(\alpha\beta\gamma\dot{\delta})(\dot{\epsilon}\dot{\zeta}\eta\theta) = (\delta\epsilon\zeta\dot{\alpha})(\dot{\beta}\dot{\gamma}\eta\theta).$$

In particular as in (19)

$$(\alpha\beta\gamma\dot{\delta})(\dot{\epsilon}\dot{\zeta})^{pq} = (\delta\epsilon\zeta\dot{\alpha})(\dot{\beta}\dot{\gamma})^{pq}.$$

Hence if we use the natural notation $(\epsilon\zeta H)$ for $\Sigma(\epsilon\zeta)^{pq} H^{rs}$ and $(\epsilon\zeta \mid G)$ for $\Sigma(\epsilon\zeta)^{pq} G_{pq}$, we have the following correlative with (21) and (24),

$$(\alpha\beta\gamma\dot{\delta})(\dot{\epsilon}\dot{\zeta}H) = (\delta\epsilon\zeta\dot{\alpha})(\dot{\beta}\dot{\gamma}H),$$

$$(\alpha\beta\gamma\dot{\delta})(\dot{\epsilon}\dot{\zeta} \mid G) = (\delta\epsilon\zeta\dot{\alpha})(\dot{\beta}\dot{\gamma} \mid G).$$

The important thing to notice is that in every detail the original identity is matched by a dual identity. They only differ in two respects:

(i) A lower suffix becomes an upper suffix, and vice versa.

(ii) Variables are replaced by formal duals, as shown by writing Greek for italic letters.

In exactly the same way any fundamental identity can be reciprocated into a dual, and there are in fact eight different modes (four direct and four dual) of writing such an identity. In all these modes the variable vectors are permuted in precisely the same way.

The case when $n = 5$ illustrates this well enough:

$$\left\{\begin{aligned}(abcd\dot{e})\,(\dot{f}\dot{g}huv) &= (\dot{a}\dot{b}efg)\,(\dot{c}\dot{d}huv)\\ (abcd\dot{e})\,(\dot{f}\dot{g})_{pq} &= (\dot{a}\dot{b}efg)\,(\dot{c}\dot{d})_{pq}\\ (abcd\dot{e})\,(\dot{f}\dot{g}\,P) &= (\dot{a}\dot{b}efg)\,(\dot{c}\dot{d}\,P)\\ (abcd\dot{e})\,(\dot{f}\dot{g}\,|\,Q) &= (\dot{a}\dot{b}efg)\,(\dot{c}\dot{d}\,|\,Q)\end{aligned}\right.$$

where $\qquad P = [P_{123}, \ldots, P_{345}],\ Q = [Q^{45}, \ldots, Q^{12}],$

are sets of $\binom{5}{\ } = \binom{5}{3} = 10$ arbitrary constants.

$$\left\{\begin{aligned}(\alpha\beta\gamma\delta\dot{\epsilon})\,(\dot{\zeta}\dot{\eta}\theta\phi\psi) &= (\dot{\alpha}\dot{\beta}\epsilon\zeta\eta)\,(\dot{\gamma}\dot{\delta}\theta\phi\psi)\\ (\alpha\beta\gamma\delta\dot{\epsilon})\,(\dot{\zeta}\dot{\eta})^{pq} &= (\dot{\alpha}\dot{\beta}\epsilon\zeta\eta)\,(\dot{\gamma}\dot{\delta})^{pq}\\ (\alpha\beta\gamma\delta\dot{\epsilon})\,(\dot{\zeta}\dot{\eta}\,R) &= (\dot{\alpha}\dot{\beta}\epsilon\zeta\eta)\,(\dot{\gamma}\dot{\delta}\,R)\\ (\alpha\beta\gamma\delta\dot{\epsilon})\,(\dot{\zeta}\dot{\eta}\,|\,S) &= (\dot{\alpha}\dot{\beta}\epsilon\zeta\eta)\,(\dot{\gamma}\dot{\delta}\,|\,S)\end{aligned}\right.$$

where $\qquad R = [R^{123}, \ldots, R^{345}],\ \ S = [S_{45}, \ldots, S_{12}]$

are sets of ten arbitrary constants.

13. The Sylvester Identity.

We conclude this investigation of the principle of duality by making a dual transformation of the Sylvester identity, §9, (II), p. 45, which was stated in the form

$$(A_i\dot{C}_k)\,(\dot{D}_iF_k) = (D_iC_k)\,(A_iF_k) \qquad i + k = n. \tag{25}$$

On the left are $\binom{k+i}{i}$ terms, due to derangement of k columns C_k and i columns D_i. Without loss of generality we can assume $i \leqslant k$.

Among these terms on the left is $(A_iC_k)\,(D_iF_k)$, while in all the rest some columns of D_i appear in the same factor as A_i.

Let us rewrite this identity with this first term transferred to the other side as

$$\sum_{q=1}\sum(A_{i+q}C_{k-q})\,(D_{i-q}F_{k+q})$$
$$=(A_iF_k)\,(D_iC_k)-(A_iC_k)\,(D_iF_k), \qquad (26)$$

where q denotes the number of columns of D_i transferred. For each value of q there are $\binom{}{q}\binom{k}{q}$ terms, because this is the number of choices which can be made from D_i and C_k. The notation A_{i+q} denotes the matrix A combined with q of the columns of D. Similarly for the other suffixes. The double summation sign is used rather than the dot notation for reasons of convenience.

Example.—If $n = 4$, $A = ab$, $D = cd$, $C = wt$, $F = uv$, then such an identity is

$$\begin{aligned}&(abuv)\,(cdwt)-(abwt)\,(cduv)\\ =\;&(abct)\,(dwuv)-(abcw)\,(dtuv)\\ -\;&(abdt)\,(cwuv)+(abdw)\,(ctuv)\\ +\;&(abcd)\,(uvwt).\end{aligned} \qquad (27)$$

On the right, there are four terms answering to $q = 1$, due to derangements of c, d and of t, w independently. We write these as $(ab\dot{c}\bar{t})\,(\dot{d}\bar{w}uv)$, with *dots* and *bars* to distinguish the two determinantal permutations. So the identity now runs

$$(abuv)\,(cdwt)-(abwt)\,(cduv)=(ab\dot{c}\bar{t})\,(\dot{d}\bar{w}uv)+(abcd)\,(uvwt). \qquad (28)$$

Since the eight columns a, b, . . . , v of these determinants are quite arbitrary, let us take u_i, v_i, w_i, t_i respectively as the co-factors of ξ_i, η_i, ζ_i, ω_i in the non-vanishing determinant $(\xi\eta\zeta\omega)$. So from the table

$$\begin{matrix} u & v & w & t \\ \xi & \eta & \zeta & \omega \end{matrix}$$

we deduce $u_1 = (\eta\zeta\omega)_{234}$, $v_1 = -(\xi\zeta\omega)_{234}$, &c. Then by the Jacobi ratio theorem $(uv)_{12} = (\xi\eta\zeta\omega)\,(\zeta\omega)_{34}$, &c. Consequently

$$(abuv) = (ab\,|\,\zeta\omega)\,(\xi\eta\zeta\omega), \quad (abct) = -(abc\,|\,\xi\eta\zeta), \text{ \&c.}$$

If these are substituted in (28) and the common factor $(\xi\eta\,\zeta\omega)^2$ is removed we obtain, as in §8, p. 89,

$$\begin{aligned}&(ab\,|\,\zeta\omega)\,(cd\,|\,\xi\eta) - (ab\,|\,\xi\eta)\,(cd\,|\,\zeta\omega)\\ &= -\,(ab\dot{c}\,|\,\xi\eta\,\zeta)\,(\dot{d}\,|\,\omega) + (ab\dot{c}\,|\,\xi\eta\omega)\,(\dot{d}\,|\,\zeta) + (abcd)\,(\xi\eta\,\zeta\omega)\\ &= -\,(ab\dot{c}\,|\,\xi\eta\,\bar{\zeta})\,(\dot{d}\,|\,\bar{\omega}) + (abcd)\,(\xi\eta\,\zeta\omega). \quad \cdot\;\cdot\;\cdot\;\cdot\;\cdot\;\cdot \quad (29)\end{aligned}$$

This is a form of the Sylvester identity in terms of matrix inner products. Since it is a polynomial identity which holds for all values of the elements concerned provided $(\xi\eta\,\zeta\omega)$ does not vanish, this last restriction may be removed as in the case of §4, p. 82. The formula therefore holds without exception.

14.[1] Formal Proof of the Sylvester Identity.

More generally, by the same methods, we may transform identity (25) to a relation between columns of four matrices A, B, P, Q of the same currency. For *if*

$$A_i = a_1 a_2 \ldots a_i, \quad B_i = b_1 b_2 \ldots b_i,$$
$$P_i = x_1 x_2 \ldots x_i, \quad Q_i = y_1 y_2 \ldots y_i,$$

then

$$\left.\begin{aligned}&\sum_{q=1}^{i} (-)^q\,(a_1 \ldots a_i \dot{b}_1 \ldots \dot{b}_q\,|\,x_1 \ldots x_i \bar{y}_1 \ldots \bar{y}_q)\\ &\qquad\qquad\qquad \times (\dot{b}_{q+1} \ldots \dot{b}_i\,|\,\bar{y}_{q+1} \ldots \bar{y}_i)\\ &= (a_1 \ldots a_i\,|\,y_1 \ldots y_i)\,(b_1 \ldots b_i\,|\,x_1 \ldots x_i)\\ &\qquad - (a_1 \ldots a_i\,|\,x_1 \ldots x_i)\,(b_1 \ldots b_i\,|\,y_1 \ldots y_i).\end{aligned}\right\} \quad (30)$$

This can be written more shortly as

$$\begin{aligned}&\sum_{q=1}^{i}\sum (A_{i+q}\,|\,P_{i+q})\,(B_{i-q}\,|\,Q_{i-q})\\ &\qquad = (A_i\,|\,P_i)\,(B_i\,|\,Q_i) - (A_i\,|\,Q_i)\,(B_i\,|\,P_i). \quad \cdot \quad (31)\end{aligned}$$

If 2i > n, *the upper summation limit is* n.

Proof.—

The result follows from (26) by the Jacobi ratio theorem, exactly as in the quaternary case just considered, provided

$2i$

---

[1] This section may be omitted on a first reading.

We take $2i+j=n$, and work with two dual matrices of order n

$$[x_1 x_2 \dots x_i y_1 y_2 \dots y_i z_1 \dots z_j]$$
$$[u_1 u_2 \dots u_i v_1 v_2 \dots v_j w_1 \dots w_j]$$

where the lower is the adjugate of the upper. For instance, the adjugate of the minor determinant

$$(x_1 \dots x_i y_1 \dots y_q)_\lambda$$

is

$$(v_{q+1} \dots v \ w_1 \dots w_j)_{\lambda'}$$

where λ, λ' are algebraic complementary suffix rows. Similarly that of $(y_{q+1} \dots . y_i)_\mu$ is $(-)^{(i-q)(i+q)} (u_1 \dots u_i v_1 \dots v_q w_1 \dots w_j)_{\mu'}$.

Hence the series on the left of (30) is equal to

$$\Sigma (-)^{q+i^2-q^2} (a_1 \dots a_i \dot{b}_1 \dots \dot{b}_q \bar{v}_{q+1} \dots \bar{v}_i w_1 \dots w_j)$$
$$\times (\dot{b}_{q+1} \dots \dot{b}_i u_1 \dots u_i \bar{v}_1 \dots \bar{v}_q w_1 \dots w_j).$$

Since q^2-q is even, the sign simplifies to $(-)^{i^2}=(-)^i$. This makes the series a Sylvester series as in §11 (26), p. 94, so that it is equal to the two terms

$$(-)^i\{(a_1 \dots a_i u_1 \dots u_i w_1 \dots w_j)$$
$$\times (b_{q+1} \dots b_i b_1 \dots b_q v_{q+1} \dots v_i v_1 \dots v_q w_1 \dots w_j)$$
$$- (a_1 \dots a_i v_1 \dots v_i w_1 \dots w_j)(b_1 \dots b_i u_1 \dots u_i w_1 \dots w_j)\}$$

which simplify to

$$(-)^i\{(a_1 \dots u_1 \dots w_1 \dots)(b_1 \dots v_1 \dots w_1 \dots)$$
$$- (a_1 \dots v_1 \dots w_1 \dots)(b_1 \dots u_1 \dots w_1 \dots)\},$$

or finally to

$$(a_1 \dots a_i \mid y_1 \dots y_i)(b_1 \dots b_i \mid x_1 \dots x_i)$$
$$- (a_1 \dots a_i \mid x_1 \dots x_i)(b_1 \dots b_i \mid y_1 \dots y_i).$$

This proves the theorem, provided $2i \leqslant n$.

To prove it if $2i > n$, we merely take the theorem for original determinants of order $2i$, where all the elements of the rows $n+1, n+2, \dots, n+(2i-n)$ are zero. For this automatically turns the inner product $\Sigma a_k x_k$ summed for $k=1, 2, \dots, 2i$

into one for $k = 1, 2, \ldots, n$, so that any compound inner product may be interpreted indifferently as of either category $2i$ or n. It further ensures that rth compound inner products when $r > n$ should vanish identically. This is needed when $n > i$, in order to remove the terms of (31) for values of q between $n - i$ and i.

Example.—If in (29), p. 95, $a_4 = b_4 = c_4 = d_4 = 0$, the quaternary identity leads to the ternary identity

$$(ab \mid \zeta\omega)(cd \mid \xi\eta) - (ab \mid \xi\eta)(cd \mid \zeta\omega) = -(ab\dot{c})(\xi\eta\bar{\zeta})(\dot{d} \mid \bar{\omega}).$$

For the third compound $(abc \mid \xi\eta\zeta)$ now resolves into factors $(abc)(\xi\eta\zeta)$.

The reader will have no difficulty in finding the dual form of the other fundamental types (I), (III) of §9, p. 45, by making analogous transformations.

EXAMPLES

1. If $n > 3$, prove

$$(bc \mid \eta\zeta)(a \mid \xi) + (ca \mid \eta\zeta)(b \mid \xi) + (ab \mid \eta\zeta)(c \mid \xi) = (abc \mid \xi\eta\zeta),$$

and adapt the result to the binary and ternary cases, $n = 2, 3$.

2. If $n > 4$, prove by dualizing ξ, &c., that

$$(bc \mid \eta\zeta)(ad \mid \xi\omega) + (ca \mid \eta\zeta)(bd \mid \xi\omega) + (ab \mid \eta\zeta)(cd \mid \xi\omega)$$
$$= (abc \mid \xi\eta\zeta)(d \mid \omega) - (abc \mid \omega\eta\zeta)(d \mid \xi).$$

3. Prove the same identity by expanding

$$\begin{vmatrix} a_\xi & b_\xi & c_\xi & d_\xi \\ a_\eta & b_\eta & c_\eta & \\ a_\zeta & b_\zeta & c_\zeta & \\ a_\omega & b_\omega & c_\omega & \end{vmatrix}$$

in two different ways and using the Binet Cauchy theorem, p. 81.

CHAPTER VI

SPECIAL TYPES OF DETERMINANT

1. Properties of Matrices and Determinants connected with the Leading Diagonal.

Associated with every square matrix A of order n is the matrix $\lambda I - A$ obtained by subtracting λ from each element of the leading diagonal, and changing all the signs. The equation in λ obtained by equating to zero the corresponding determinant $(\lambda I - A)$ is called the *characteristic equation* of the matrix A.

For example, if $n = 3$, and $A = [a_{ij}]$,

$$f(\lambda) = |\lambda I - A| = \begin{vmatrix} \lambda - a_{11} & -a_{12} & -a_{13} \\ -a_{21} & \lambda - a_{22} & -a_{23} \\ -a_{31} & -a_{32} & \lambda - a_{33} \end{vmatrix} = 0.$$

This equation is a cubic in λ, and in general the characteristic equation is of order n, as is apparent by writing down the leading term in the expansion of this determinant. So

$$f(\lambda) = a_0 \lambda^n + a_1 \lambda^{n-1} + \ldots + a_i \lambda^{n-i} + \ldots + a_n, \quad a_0 = 1,$$

where each a_i is a polynomial function of the n^2 elements a_{ij}. The n roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of this equation $f(\lambda) = 0$ are called the *latent roots* of the matrix.

As it is useful to know how to perform the expansion of this characteristic determinant, a method is suggested by the following examples.

EXAMPLES

1. Prove

$$\begin{array}{cccc} a_1 + x & b_1 & c_1 & d_1 \\ & b_2 + x & c_2 & \\ & b_3 & c_3 + x & d_3 \\ & b_4 & & d_4 + x \end{array}$$

$$= x^4 + (a_1 + b_2 + c_3 + d_4)x^3 + \{ |b_2 c_3| + |a_1 d_4| + |a_1 c_3| + |b_2 d_4|$$
$$+ |a_1 b_2| + |c_3 d_4| \} x^2 + \{ |b_2 c_3 d_4| + |a_1 c_3 d_4| + |a_1 b_2 d_4|$$
$$+ |a_1 b_2 c_3| \} x + |a_1 b_2 c_3 d_4|.$$

[Differentiate both sides of the identity with regard to x four times. Put $x = 0$ at each stage.]

2. Generalize this result.

3. If $p_1, p_2, \ldots, p_n$ are the leading diagonal elements of a determinant Δ, and $P, P_i, P_{ij}, P_{ijk}, \ldots$ denote the values (after putting $p_1 = p_2 = \ldots = 0$) of Δ, the co-factor of p_i, that of $p_i p_j$, &c., show that Δ can be expressed in the form

$$P + \Sigma P_i p_i + \Sigma P_{ij} p_i p_j + \ldots + p_1 p_2 \ldots p_n.$$

4. Prove

$$\begin{vmatrix} p_1 & b_1 & c_1 \\ a_2 & p_2 & c_2 \\ a_3 & b_3 & p_3 \end{vmatrix} = \begin{vmatrix} . & b_1 & c_1 \\ a_2 & . & c_2 \\ a_3 & b_3 & . \end{vmatrix} + \begin{vmatrix} . & c_2 \\ b_3 & . \end{vmatrix} p_1$$

$$+ \begin{vmatrix} . & c_1 \\ a_3 & . \end{vmatrix} p_2 + \begin{vmatrix} . & b_1 \\ a_2 & . \end{vmatrix} p_3 + p_1 p_2 p_3.$$

2. The Cayley Hamilton Theorem.

For a second order matrix the characteristic equation is a quadratic. Thus if $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$,

$$f(\lambda) = \begin{vmatrix} \lambda - a & -b \\ -c & \lambda - d \end{vmatrix} = \lambda^2 - (a + d)\lambda + ad - bc = 0.$$

If we construct the corresponding function $f(A)$ of the matrix A itself by evaluating

$$A^2 - (a + d)A + (ad - bc)I$$

as a second order matrix, we obtain the remarkable result that all the elements of this matrix are zero. This can readily be verified. It is an instance of an important theorem which runs as follows.

THE CAYLEY HAMILTON THEOREM.—*Every square matrix satisfies its own characteristic equation.*

Proof.—

Let the matrix $\lambda I - A$, constructed from a given matrix A of order n, have for its adjugate (§8, p. 70) B. Since the elements of $\lambda I - A$ are at most of the first degree in λ, their co-factors in $| \lambda I - A |$ are at most of degree $n - 1$ in λ. Hence we may write the typical element β_{ij} of the adjugate as

$$b_0 + b_1 \lambda + \ldots + b_{n-1} \lambda^{n-1},$$

where the n coefficients b_i are polynomial functions of the n^2 elements of A.

Thus the matrix B itself may be written as

$$B_0 + B_1 \lambda + \ldots + B_{n-1} \lambda^{n-1},$$

where B_k is a matrix ~~whose typical element~~ is b_k

But since $|\lambda I - A| \equiv f(\lambda)$ is the determinant of $\lambda I - A$, it follows by direct multiplication (§8, (11), p. 70) that

$$(\lambda I - A)B = |\lambda I - A| \, I = f(\lambda) I.$$

Hence

$$\lambda B - AB = f(\lambda) I,$$

or

$$\lambda \sum_{k=0}^{n-1} B_k \lambda^k - A \sum_{k=0}^{n-1} B_k \lambda^k = f(\lambda) I$$

$$= (a_0 \lambda^n + a_1 \lambda^{n-1} + \ldots + a_n) I.$$

This is an identity for all values of λ. Equating coefficients of $\lambda^n, \lambda^{n-1}, \ldots, \lambda^1, \lambda^0$ we obtain

$$\begin{aligned}
B_{n-1} \qquad\qquad &= a_0 I \\
B_{n-2} \quad - AB_{n-1} &= a_1 I \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad &\cdot \quad \cdot \quad \cdot \\
B_{n-k-1} - AB_{n-k} &= a_k I \\
\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad &\cdot \quad \cdot \quad \cdot \\
- AB_0 \quad &= a_n I.
\end{aligned}$$

By fore multiplication with $A^n, A^{n-1}, \ldots, A, 1$ respectively and addition, we obtain

$$0 = a_0 A^n + a_1 A^{n-1} + \ldots + a_{n-1} A + a_n I,$$

which proves the theorem.

It is to be noted that the theorem holds for singular matrices.

There are alternative ways of writing this result. Since we can factorize the expression $f(\lambda)$ in terms of its latent roots λ_i as

$$f(\lambda) = a_0 (\lambda - \lambda_1)(\lambda - \lambda_2) \ldots (\lambda - \lambda_n),$$

we can therefore also write

$$f(A) = a_0 (A - \lambda_1 I)(A - \lambda_2 I) \ldots (A - \lambda_n I),$$

where the order of factors is immaterial (cf. Ex. **10**, p. 72). Here is a case where, by the Cayley Hamilton theorem, the product of n matrices $A - \lambda_i I$ is zero, though we cannot assume that any of the factors are zero.

EXAMPLES

1. If Λ is the diagonal matrix $\begin{bmatrix} \lambda_1 & \cdot & \cdot \\ \cdot & \lambda_2 & \cdot \\ \cdot & \cdot & \lambda_3 \end{bmatrix}$ prove that Λ also satisfies the characteristic equation $f(\lambda) = 0$ of the matrix A.

2. If B is an arbitrary non-singular matrix, prove that $B\Lambda B^{-1}$ and $B^{-1}\Lambda B$ both satisfy $f(\lambda) = 0$.

3. Show that this is true for a square matrix of any order.

4. Prove the latent roots of the reciprocal of the third order matrix A are $\lambda_1{}^{-1}$, $\lambda_2{}^{-1}$, $\lambda_3{}^{-1}$.

5. If the latent roots of A are λ, λ, μ, prove that $f(\lambda) = 0$ is satisfied by $\Lambda = \begin{bmatrix} \lambda & 1 & \cdot \\ \cdot & \lambda & \cdot \\ \cdot & \cdot & \mu \end{bmatrix}$.

6. Verify that $\Lambda = \begin{bmatrix} \lambda & 1 & \cdot \\ \cdot & \lambda & 1 \\ \cdot & \cdot & \lambda \end{bmatrix}$ satisfies $f(\lambda) = 0$ if the three roots are ıal.

7. Show that $B\Lambda B^{-1}$ and $B^{-1}\Lambda B$ also satisfy $f(\lambda) = 0$, in **6**.

3. Special Types of Determinant.

Bordered determinants. Symmetric and skew symmetric determinants.

If above and to the left of a square matrix $[a_{ij}]$ of order n we add a row and column

$$\begin{array}{ccc} 0 & u_1 & \qquad\qquad u_n \\ v_1 & & \end{array}$$

we obtain a bordered matrix of order $n + 1$. Its determinant is also said to be bordered and is written shortly as

$$\begin{vmatrix} \cdot & u_i \\ v_i & a_{ij} \end{vmatrix}.$$

So if $n = 3$ an example is

$$\Sigma_1 = \begin{array}{cccc} \cdot & u_1 & u_2 & u_3 \\ v_1 & a_1 & b_1 & c_1 \\ v_2 & a_2 & b_2 & c_2 \\ v_3 & a_3 & b_3 & c_3 \end{array}$$

We may border more deeply by adding a double row and column meeting in a set of four zeros. We should write

$$\Sigma_2 = \begin{array}{ccccc} & \cdot & u_1 & u_2 & u_3 \\ & \cdot & u_1' & u_2' & u_3' \\ v_1 & v_1' & a_1 & b_1 & c_1 \\ v_2 & v_2' & a_2 & '_2 & c_2 \\ v_3 & v_3' & a_3 & b_3 & \end{array}$$

Such a process can be generalized, giving what may be called bordered determinants of the first, second, . . . rth orders derived from a nucleus Δ. Hence if θ_r, ϕ_r each denote matrices of n rows and r columns, or briefly matrices of currency r, and if as usual the accent indicates transposition, then we can write the general bordered determinant derived from

$$\Delta = | a_{ij} |$$

$$\Sigma_r = \qquad a_{ij}$$

Now consider values of r between 0 and n, of which the above case Σ_2 is typical. Expand Σ_r by the $\left(\begin{matrix} r \\ \cdots\cdots \\ n \end{matrix}\right)$ Laplace development. For Σ_2, $(n = 3)$, the result is three terms

$$(uu')_{23}\,(vv'a)_{123} + (uu')_{31}\,(vv'b)_{123} + (uu')_{12}\,(vv'c)_{123}.$$

Expand each factor $(vv'a)$ by the $(r \,\vdots\, n - r)$ development and the result is linear in the set

$$(vv')_{23}, \quad (vv')_{31}, \quad (vv')_{12}.$$

We infer that Σ_r is *bilinear* in the two sets of determinants of the border matrices θ'_r, ϕ_r.

If $r = n$ the same argument shows that the bordered determinant Σ_n is merely the product

$$(-)^n\,(uu' \ldots)\,(vv' \ldots),$$

better written

$$(-)^n\,| u_{ij} | \;\; | v_{ij} |.$$

If $r = 0$, Σ_0 is Δ. If $r > n$, $\Sigma_r = 0$.

4. Reciprocation of Bordered Determinants.

Bordered determinants obey the principle of duality in a manner which recalls the Jacobi ratio theorem. In fact the following results may be regarded as a corollary of this theorem. For consider three square matrices of order n

$$A = \begin{bmatrix} a_1 & b_1 & \dots & m_1 \\ a_2 & b_2 & \dots & m_2 \\ \\ a_n & b_n & \dots & m_n \end{bmatrix}, \quad X = \begin{bmatrix} x_1 & y_1 & \dots & t_1 \\ x_2 & y_2 & \dots & t_2 \\ \cdot & \cdot & \cdot & \cdot \\ x_n & y_n & \dots & t_n \end{bmatrix}, \quad \Xi = \begin{bmatrix} \xi_1 & \eta_1 & \dots & \omega_1 \\ \xi_2 & \eta_2 & \dots & \omega_2 \\ \cdot & \cdot & \cdot & \cdot \\ \xi_n & \eta_n & \dots & \omega_n \end{bmatrix}$$

together with their reciprocals with raised suffixes. For brevity take $n = 3$, so that, as before, $|A|\, a^1 = |b_2 c_3|$, &c., while $|X|\, x^1 = |y_2 z_3|$, &c., and $|\Xi|\, \xi^1 = |\eta_2 \zeta_3|$, &c. Then the following identities will hold:

$$\begin{matrix} \cdot & \cdot & \cdot & x_1 & x_2 & x_3 \\ \cdot & \cdot & \cdot & y_1 & y_2 & y_3 \\ \cdot & \cdot & \cdot & z_1 & z_2 & z_3 \\ \xi_1 & \eta_1 & \zeta_1 & a_1 & b_1 & c_1 \\ \xi_2 & \eta_2 & \zeta_2 & a_2 & b_2 & c_2 \\ \xi_3 & \eta_3 & \zeta_3 & a_3 & b_3 & c_3 \end{matrix} = \rho \begin{matrix} a^1 & b^1 & c^1 \\ a^2 & b^2 & c^2 \\ a^3 & b^3 & c^3 \end{matrix}$$

$$\begin{matrix} \cdot & \cdot & y_1 & y_2 & y_3 \\ \cdot & \cdot & z_1 & z_2 & z_3 \\ \eta_1 & \zeta_1 & a_1 & b_1 & c_1 \\ \eta_2 & \zeta_2 & a_2 & b_2 & c_2 \\ \eta_3 & \zeta_3 & a_3 & b_3 & c_3 \end{matrix} = \rho \begin{matrix} \cdot & x^1 & x^2 & x^3 \\ \xi^1 & a^1 & b^1 & c^1 \\ \xi^2 & a^2 & b^2 & c^2 \\ \xi^3 & a^3 & b^3 & c^3 \end{matrix}$$

$$\begin{vmatrix} \cdot & z_1 & z_2 & z_3 \\ \zeta_1 & a_1 & b_1 & c_1 \\ \zeta_2 & a_2 & b_2 & c_2 \\ \zeta_3 & a_3 & b_3 & c_3 \end{vmatrix} = \rho \begin{vmatrix} \cdot & \cdot & x^1 & x^2 & x^3 \\ \cdot & \cdot & y^1 & y^2 & y^3 \\ \xi^1 & \eta^1 & a^1 & b^1 & c^1 \\ \xi^2 & \eta^2 & a^2 & b^2 & c^2 \\ \xi^3 & \eta^3 & a^3 & b^3 & c^3 \end{vmatrix}$$

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix} = \rho \begin{vmatrix} \cdot & \cdot & \cdot & x^1 & x^2 & x^3 \\ \cdot & \cdot & \cdot & y^1 & y^2 & y^3 \\ \cdot & \cdot & \cdot & z^1 & z^2 & z^3 \\ \xi^1 & \eta^1 & \zeta^1 & a^1 & b^1 & c^1 \\ \xi^2 & \eta^2 & \zeta^2 & a^2 & b^2 & c^2 \\ \xi^3 & \eta^3 & \zeta^3 & a^3 & b^3 & c^3 \end{vmatrix}$$

where $\rho = -\,|a_1 b_2 c_3|\;|x_1 y_2 z_3|\;|\xi_1 \eta_2 \zeta_3|$.

The proof is immediate, by expanding the left-hand side determinants as bilinear functions of the border matrices, and then raising the suffixes by use of the Jacobi ratio theorem.

For instance, in the third identity, a typical term involving $z_2 \zeta_1$ is $z_2 \zeta_1 \mid a_2 c_3 \mid$. But by the ratio theorem

$$z_2 = \mid x_1 y_2 z_3 \mid \quad \mid x^3 y^1 \mid, \qquad \zeta_1 = \mid \xi_1 \eta_2 \zeta_3 \mid \quad \mid \xi^2 \eta^3 \mid$$

$$\mid a_2 c_3 \mid = - \mid a_1 b_2 c_3 \mid b^1,$$

whence

$$z_2 \zeta_1 \mid a_2 c_3 \mid = \rho \mid x^3 y^1 \mid \quad \mid \xi^2 \eta^3 \mid \quad b^1,$$

agreeing with the corresponding term on the right.

In general for n-rowed matrices A, X, Ξ we have

$$\rho = (-)^n \mid A \mid \quad \mid X \mid \quad \mid \Xi \mid,$$

and the letters absent in the borders of determinants on one side of the identities are present on the other, their arrangement being determined by the algebraic complement rule as in Jacobi's theorem.

5. Bordered Adjugate Determinant.

As in the Jacobi theorem itself it is useful to recognize the earlier form of the theorems of the last paragraph. Namely, when all elements with upper suffixes are replaced by capital letters denoting co-factors of elements with lower suffixes, the theorems hold if ρ is multiplied by suitable positive integral powers of $\mid a_1 b_2 c_3 \mid$.

6. Symmetrical Matrices and Determinants.

These are symmetrical if transposition of rows into columns makes no difference, so that

$$a_{ij} = a_{ji} \qquad \{ij\}.$$

For example,

$$\begin{vmatrix} a & h \\ h & b \end{vmatrix}, \quad \begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}, \quad \begin{vmatrix} a & h & g & u \\ h & b & f & v \\ g & f & c & w \\ u & v & w & d \end{vmatrix}$$

The condition that a matrix A should be symmetrical can be written $A = A'$.

EXAMPLES

1. If capital letters denote co-factors of corresponding small letters in Δ: $\begin{vmatrix} a & h & g \\ h & b & f \\ g & f & c \end{vmatrix}$, prove $\begin{vmatrix} \cdot & u & v & w \\ u & a & h & g \\ v & h & b & f \\ w & g & f & c \end{vmatrix}$

$$= -(Au^2 + Bv^2 + Cu^2 + 2Fvw + 2Gwu + 2Huv),$$

$$\begin{matrix} \cdot & \cdot & u & v & w \\ \cdot & \cdot & u^1 & v^1 & w^1 \\ u & u^1 & a & h & g \\ v & v^1 & h & b & f \\ w & w^1 & g & f & c \end{matrix} = ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$$

where $x, y, z = \left\| \begin{matrix} u & v & w \\ u^1 & v^1 & w^1 \end{matrix} \right|$

Expand as a quadratic in the same way

$$\begin{matrix} \cdot & x & y & z \\ x & A & H & G \\ y & H & B & F \\ z & G & F & C \end{matrix}$$

Answer: $-\Delta\,(ax^2 + by^2 + \&\text{c.})$.

2. If A is an arbitrary square or rectangular matrix, and A' is its transposed, then AA' is symmetrical, and so is $A'A$.

[Use the fundamental relation of type $(BC)' = C'B'$.]

7. Skew Symmetric Determinants.

$\Delta = |\,a_{ij}\,|$ is skew symmetric if $a_{ij} = -a_{ji}$, so that on the leading diagonal every element is zero, since

$$a_{ii} = -a_{ii} = 0.$$

In this case the matrix $[a_{ij}]$ is said to alternate in its double suffixes: interchange of suffixes is accompanied by change of sign. So interchange of the set of rows and the set of columns is accompanied by n changes of sign, one for each row, n being the order of the determinant. Thus if A is the matrix, and A' its transposed,

$$A = -A', \quad \text{but } |\,A\,| = (-)^n\,|\,A'\,|,$$

whence $\Delta = (-)^n\Delta$. Accordingly we have the theorem:

If n *is odd,* Δ *is identically zero.*

A skew symmetric matrix or determinant is completely specified by the $\frac{1}{2}n(n-1)$ elements in the triangle above its diagonal. Thus if $n = 3$, we might have

$$0 = \Delta = \begin{vmatrix} \cdot & c & b \\ -c & \cdot & a \\ -b & -a & \cdot \end{vmatrix}$$

$$A = \begin{bmatrix} \cdot & c & b \\ -c & \cdot & a \\ -b & -a & \cdot \end{bmatrix}, \quad A' = \begin{bmatrix} \cdot & -c & -b \\ c & \cdot & -a \\ b & a & \cdot \end{bmatrix}$$

If, however, n *is even,* Δ *is a perfect square function of its elements.*

For let

$$\Delta_n = \begin{vmatrix} \cdot & a & b & c & \dots \\ -a & \cdot & d & e & \dots \\ -b & -d & \cdot & f & \dots \\ -c & -e & -f & \cdot & \dots \end{vmatrix}$$

Consider the co-factors of the leading four elements $\begin{matrix} \cdot & a \\ -a & \cdot \end{matrix}$. Since the diagonal co-factors are manifestly skew symmetric of order $n-1$, which is odd, they vanish. Also if A is co-factor of a, that of $-a$ is the transposed of A with every sign changed. Hence it is $(-)^{n-1}A, = -A$. The determinant of these four co-factors is therefore

$$\begin{vmatrix} 0 & A \\ -A & 0 \end{vmatrix}$$

But by the Jacobi ratio theorem this can be written

$$\begin{vmatrix} \cdot & f & \dots \\ -f & \cdot & \dots \\ \cdot & \cdot & \cdot & \cdot \end{vmatrix} = \Delta_n \Delta_{n-2}$$

say, where Δ_{n-2} is also an even skew symmetric determinant. Thus if Δ_{n-2} is a perfect square, so is Δ_n. But Δ_2 obviously is so. Hence by induction so is Δ_n, or else it is zero. It is not zero in general, since in the special case when $b = c = \dots$ all vanish except $a, f \dots$, $\Delta = a^2 f^2 \dots$, where $a, f, \dots$ are the letters occurring alternately in the positions nearest the leading diagonal.

8. Characteristic Function of a Skew Matrix. (Weierstrass, 1879.)

Let a determinant be expanded by its principal diagonal as in § 1. In particular if the determinant

$$\begin{vmatrix} \lambda & c & -b \\ -c & \lambda & a \\ b & -a & \lambda \end{vmatrix}$$

is so expanded, the result is

$$(a^2 + b^2 + c^2)\lambda + \lambda^3.$$

Suppose S is a skew symmetric matrix and $|\lambda I + S|$ is the corresponding determinant with λ replacing each zero in the leading diagonal, as in the above example for the third order case. Expanding by its leading diagonal in an ascending series for λ we have

$$|\lambda I + S| = P + \Sigma P_i \lambda + \Sigma P_{ij} \lambda^2 + \ldots + \lambda^n$$

where P_i is the co-factor of the ith diagonal element in $|S|$, and P_{ij} that of the product of the ith and jth diagonal elements in $|S|$, and so on. But all such co-factors are skew symmetric; hence those of odd order vanish, and those of even order are perfect squares. Thus reversing the terms of the series, we have

$$|\lambda I + S| = \lambda^n + Q\lambda^{n-2} + R\lambda^{n-4} + \ldots,$$

where Q, R, . . . are sums of squares and therefore are essentially positive if the elements a, b, c, . . . of S are real.

Hence the matrix $\lambda I + S$ cannot be singular if its elements are real, as long as $\lambda > 0$. In particular $I \pm S$ gives two non-singular matrices.

9. Summary of Theorems on Compound Determinants.

In spite of the great intrinsic interest of the subject, and the wonderful flexibility of determinants as practical working tools in many branches of pure and applied mathematics, there is still a considerable absence of systematic knowledge of even the main results in the theory. It may therefore be of help to the reader to have a short statement of at any rate one main branch of what is indeed a very wide subject.

We may sum up[1] the theory of compound determinants in eight related theorems. These appear in their relative positions most clearly if a numerical notation is adopted, where digits have the significance of letters in what has preceded, and in addition a group of less than *n* digits indicates a certain minor.

$$\text{I.}\quad ((234)(134)(124)(123)) = (1234)^3.$$

Cauchy's theorem on the adjugate, 1812: *The adjugate is the* (n — 1)*th power of the original determinant* (§7, p. 67).

$$\text{II.}\quad ((134)(124)(123)) = (1234)^2(1).$$

Jacobi's theorem on the adjugate, 1831: *A minor of order* r *of the adjugate is equal to the complementary minor in the original determinant* A *multiplied by the* (r — 1)*th power of* A.

$$\text{III.}\quad ((12)(13)(14)(23)(24)(34)) = (1234)^3.$$

Sylvester's theorem on the *m*th compound, 1851: *The* m*th compound of a given determinant* A *is the* $\binom{n-1}{m-1}$*th power of* A. (§9, p. 87).

$$\text{IV.}\quad ((14)(23)(24)(34)) = (1234)((34)(24)).$$

Franke's theorem on the *m*th compound, 1862: *A minor of order* r *of the* m*th compound is equal to the complementary minor in the adjugate* (n — m)*th compound multiplied by the* $\left(r-\binom{n-1}{m}\right)$*th power of the original determinant.*

$$\text{V.}\quad ((a23)(1b3)(12c)) = (123)^2(abc).$$

Bazin's theorem, 1851: *If the determinants obtained by replacing a column of* A *in all possible ways by a column of* B *are elements of a "hybrid" compound determinant, the latter is equal to* $A^{n-1}B$ (p. 56, Ex. 8).

$$\text{VI.}\quad ((1b34)(12c4)(123d)) = (1234)^2(1bcd).$$

Reiss' theorem, 1867: *A minor of order* r *in the Bazin hybrid compound of* A *and* B *is equal to the complementary minor in the*

---

[1] I owe the following illuminating summary to Dr. A. C. Aitken. Cf. his *Determinants and Matrices* (Edinburgh, 1939), pp. 101–103.

reciprocal hybrid (*i.e. that in which the rôles of* A *and* B *are interchanged*) *multiplied by a power* A^{r-1} *of* A, *and a power* B^{r-n+1} *of* B.

VII. $((ab34)\,(a2c4)\,(a23d)\,(1bc4)\,(1b3d)\,(12cd)) = (1234)^3\,(abcd)^3$.

Reiss' theorem, 1867 (Picquet, 1878): *The hybrid compound of* A *and* B *whose elements are the determinants obtained by replacing in all possible ways* m *columns of* A *by* m *columns of* B *is equal to* $A^{\binom{n-1}{m}}\,B^{\binom{n-1}{m-1}}$.

Bazin's theorem is the case $m = 1$.

VIII. $((ab34)\,(a2c4)\,(a23d)\,(1bc4)) = (1234)\,(abcd)\,((a2c4)\,(ab34))$.

Reiss' theorem, 1867 (Picquet, 1878): *A minor of order* r *in the Reiss hybrid compound is equal to the complementary minor in the reciprocal Reiss hybrid, multiplied by*

$$A^{r-\binom{n-1}{m-1}}\;B^{r-\binom{n-1}{m}}.$$

Theorem VI is the case when $m = 1$.

The above are the eight chief results in their actual order of discovery. Theorems I, II, III, V alone have been proved in the preceding pages, but the others can be dealt with by the same methods.

Theorem II in the notation of p. 52 would follow from

$$(134\,.\,124\,.\,123\,.\,pqr) = (1234)^2\,(pqr1)$$

by decomposing the first matrix 134. Jacobi's theorem then follows by equating the various coefficients of $(pqr)_{ijk}$ on both sides of this identity. This in fact gives another proof for what has been called the Jacobi ratio theorem in §3, p. 78.

We may very properly write

$$\mathrm{I} : \mathrm{II} :: \mathrm{III} : \mathrm{IV}$$

to show the relation of the first four of these theorems.

CHAPTER VII

DIFFERENTIATION OF A DETERMINANT

1. The Polarizing Process.

When the general n-rowed determinant

$$\Delta \equiv | e_{ij} | \equiv | a_1 b_2 c_3 \ldots m_n |$$

is regarded as a function of its n^2 different elements, treated as independent variables, it yields the result $\dfrac{\partial \Delta}{\partial e_{ij}} = E_{ij}$, where E_{ij} is the co-factor of e_{ij}. This is simply because Δ is a linear function in the single quantity e_{ij}.

Again, since

$$\Delta = a_1 A_1 + a_2 A_2 + \ldots + a_n A_n \quad . \quad . \qquad (1)$$

and $A_i = \partial \Delta \,|\, \partial a_i$, it follows that

$$\Delta = a_1 \frac{\partial \Delta}{\partial a_1} + a_2 \frac{\partial \Delta}{\partial a_2} + \ldots + a_n \frac{\partial \Delta}{\partial a_n} \qquad (2)$$

which may be abbreviated to

$$\Delta = \left(a \,\middle|\, \frac{\partial \Delta}{\partial a} \right) = \left(a \,\middle|\, \frac{\partial}{\partial a} \right) \Delta \qquad (3)$$

the latter introducing the notation which separates the differential operator from its operand. In such a case it must clearly be noted that a and $\dfrac{\partial}{\partial a}$ are not commutative.

For instance, if $n = 2$,

$$\left(a \quad \frac{\partial}{\partial a} \right) \Delta = a_1 \frac{\partial \Delta}{\partial a_1} + a_2 \frac{\partial \Delta}{\partial a_2},$$

whereas

$$\begin{aligned} \left(\frac{\partial}{\partial a} \quad a \right) \Delta &= \frac{\partial}{\partial a_1} a_1 \Delta + \frac{\partial}{\partial a_2} a_2 \Delta \\ &= \Delta \frac{\partial a_1}{\partial a_1} + a_1 \frac{\partial \Delta}{\partial a_1} + \Delta \frac{\partial a_2}{\partial a_2} + a_2 \frac{\partial \Delta}{\partial a_2} \\ &= 3\Delta. \end{aligned} \qquad (4)$$

The identity (2) is a particular case of Euler's theorem for a homogeneous function of degree s in its variables. Thus if $f(x_1, x_2, \ldots, x_n)$ is such a function

$$\frac{\partial}{\partial x}\Big) f \equiv x_1 \frac{\partial f}{\partial x_1} + \ldots + x_n \frac{\partial f}{\partial x_n} = sf. \qquad (5)$$

The determinant may equally well be differentiated following any row or column. Accordingly in the double suffix notation we have results analogous to (2), such as

$$\Delta = e_{1j} \frac{\partial \Delta}{\partial e_{1j}} + e_{2j} \frac{\partial \Delta}{\partial e_{2j}} + \quad . + e_{nj} \frac{\partial \Delta}{\partial e_{nj}}. \qquad (6)$$

$$\Delta = e_{i1} \frac{\partial \Delta}{\partial e_{i1}} + e_{i2} \frac{\partial \Delta}{\partial e_{i2}} + \quad . + e_{in} \frac{\partial \Delta}{\partial e_{in}} \qquad (7)$$

Again, since

$$0 = b_1 A_1 + b_2 A_2 + \ldots + b_n A_n$$

therefore

$$0 = b_1 \frac{\partial \Delta}{\partial a_1} + \ldots + b_n \frac{\partial \Delta}{\partial a_n} = \qquad \frac{\partial}{\partial a}\Big) \Delta, \qquad (8)$$

and likewise for any other such pairs of columns, other than the a and b column used here.

More generally, if $\{x_1, x_2, \ldots, x_n\}$ denote the n elements of an arbitrary column x, not necessarily contained in Δ, *the operator*

$$\left(x \middle| \frac{\partial}{\partial a}\right) \equiv x_1 \frac{\partial}{\partial a_1} + \ldots + x_n \frac{\partial}{\partial a_n}. \quad . \quad . \quad . \qquad (9)$$

has the effect on Δ *of substituting the column of* x*'s for that of* a*'s.*

Similar remarks apply to the rows.

Hence the effect of altering a determinant by substituting a new column or a new row for an original column or row is attained by a differential operator: and this operator, as in (9), is of the inner product type. Such a process, which is very common both in algebraic geometry and in the theory of invariants, is called a *polarizing* process.

Various notation has been used for this process, acting upon a function f of variables $x_1, x_2, \ldots, x_n$, such as

$$\left(y \frac{\partial}{\partial x}\right) \equiv \left(y \quad \frac{\partial}{\partial x}\right) \equiv \sum_{i=1}^{n} y_i \frac{\partial}{\partial x_i} \equiv D_{xy} \equiv y_x. \quad . \qquad (10)$$

2. The Capelli Operators.

The process may be repeated, with the use of several sets $\{x_1, \ldots, x_n\}$, $\{y_1, \ldots, y_n\}$, ... to replace columns of the determinant. Thus if $\Delta = (abc \ldots m)$,

$$\left.\begin{aligned} \left(x \,\middle|\, \frac{\partial}{\partial a}\right)\Delta &= (xbc \ldots m), \\ \left(y \,\middle|\, \frac{\partial}{\partial b}\right)\left(x \,\middle|\, \frac{\partial}{\partial a}\right)\Delta &= (xyc \ldots m), \end{aligned}\right\} \quad \ldots \quad (11)$$

and so on.

Let us now suppose that all these sets x, y, ... are perfectly arbitrary but independent of a and b, when we regard all the elements a_i, b_j, ... as variables, so that the ordinary laws of scalar numbers may apply to expressions involving x_i, y_j, $\frac{\partial}{\partial a_k}$, &c. It follows that in the above result

$$\left(y \,\middle|\, \frac{\partial}{\partial b}\right)\left(x \,\middle|\, \frac{\partial}{\partial a}\right) \text{ is equivalent to } \left(x \,\middle|\, \frac{\partial}{\partial a}\right)\left(y \,\middle|\, \frac{\partial}{\partial b}\right), \quad (12)$$

for the x standing to the right of $\frac{\partial}{\partial b}$ is unaffected by the differentiation. This is a feature which is probably familiar to the reader through the study of linear differential equations with constant coefficients.

Next let the set x be interchanged with the set y, to give a new identity

$$\left(x \,\middle|\, \frac{\partial}{\partial b}\right)\left(y \,\middle|\, \frac{\partial}{\partial a}\right)\Delta = (yxc \ldots m). \quad \ldots \quad (13)$$

Here on the right we may write $-(xyc \ldots m)$ by interchanging two columns of the determinant. On subtracting (13) from (11)

$$\begin{array}{cc} \left(x \,\middle|\, \frac{\partial}{\partial a}\right), & \left(y \,\middle|\, \frac{\partial}{\partial a}\right) \\ \left(x \,\middle|\, \frac{\partial}{\partial b}\right), & \left(y \,\middle|\, \frac{\partial}{\partial b}\right) \end{array} \Delta = 2(xyc \ldots m) \qquad (14)$$

and by the theorem of corresponding matrices (§4, p. 79) the left-hand side is naturally written

$$\Sigma\,(xy)_{ij}\left(\frac{\partial^2}{\partial a_i\,\partial b_j} - \frac{\partial^2}{\partial a_j\,\partial b_i}\right)\Delta = \left(xy \,\middle|\, \frac{\partial}{\partial a}\;\frac{\partial}{\partial b}\right)\Delta$$

$$= -\left(xy \,\middle|\, \frac{\partial}{\partial b}\;\frac{\partial}{\partial a}\right)\Delta = \&\text{c.} \qquad (15)$$

The notation must therefore be used with caution, for in the operators the symbol $\frac{\partial}{\partial a}\,\frac{\partial}{\partial b}$ is not short for $\frac{\partial^2}{\partial a\,\partial b}$ but for various determinantal expressions; and it alternates in a and b as it is a matrix inner product (§5, p. 82).

Proceeding in this way until n auxiliary sets $x, y, z, \ldots, t$ are involved, we obtain the following identities, which for shortness are written out for the case when $n = 4$,

$$\begin{aligned}
\Delta &= (abcd) = |\, a_1 b_2 c_3 d_4 \,| \\
\left(x \,\middle|\, \frac{\partial}{\partial a}\right)\Delta &= (xbcd), \\
\left(xy \,\middle|\, \frac{\partial}{\partial a}\;\frac{\partial}{\partial b}\right)\Delta &= 2\,!\,(xycd), \\
\left(xyz \,\middle|\, \frac{\partial}{\partial a}\;\frac{\partial}{\partial b}\;\frac{\partial}{\partial c}\right)\Delta &= 3\,!\,(xyzd), \\
(xyzt)\left(\frac{\partial}{\partial a}\;\frac{\partial}{\partial b}\;\frac{\partial}{\partial c}\;\frac{\partial}{\partial d}\right)\Delta &= 4\,!\,(xyzt).
\end{aligned} \qquad (16)$$

These operators are sometimes known as the Capelli operators, while the last of the series introduces us to the very important special case of such, involving the Cayley Ω operator constructed from n^2 independent variables:

$$\Omega \equiv \begin{vmatrix} \frac{\partial}{\partial a_1} & \frac{\partial}{\partial a_2} & & \frac{\partial}{\partial a_n} \\ \frac{\partial}{\partial b_1} & \frac{\partial}{\partial b_2} & & \frac{\partial}{\partial b_n} \\ & & & \cdot\;\cdot \\ \frac{}{\partial m_1} & \frac{}{\partial m_2} & & \frac{\partial}{\partial m_n} \end{vmatrix}$$

3. The Cayley Operator.

THEOREM.—*The effect of the Cayley operator upon the* s*th power of its determinant* Δ *is* $\mathrm{s}(\mathrm{s}+1)\ldots(\mathrm{s}+\mathrm{n}-1)\Delta^{\mathrm{s}-1}$.

It will be noted that if $n=1$ this reverts to the well-known $da^s/da = sa^{s-1}$. So the theorem gives a very interesting generalization of an elementary fact.

For clearness we consider the proof[1] when $n=3$, and

$$\Omega = \left.\begin{matrix} \dfrac{\partial}{\partial a_1} & \dfrac{\partial}{\partial a_2} & \dfrac{\partial}{\partial a_3} \\ \dfrac{\partial}{\partial b_1} & \dfrac{\partial}{\partial b_2} & \dfrac{\partial}{\partial b_3} \\ \dfrac{\partial}{\partial c_1} & \dfrac{\partial}{\partial c_2} & \dfrac{\partial}{\partial c_3} \end{matrix}\right|, \quad \Delta = \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

Since $\partial\Delta/\partial a_i = A_i$, we have

$$\partial\Delta^s/\partial a_i = s\Delta^{s-1}(b_j c_k - b_k c_j), \quad (i, j, k = 1, 2, 3):$$

whence

$$\left(x_1\frac{\partial}{\partial a_1} + x_2\frac{\partial}{\partial a_2} + x_3\frac{\partial}{\partial a_3}\right)\Delta^s = s\Delta^{s-1}(x_1 A_1 + x_2 A_2 + x_3 A_3)$$

or

$$\left(x\,\middle|\,\frac{\partial}{\partial a}\right)\Delta^s = s\Delta^{s-1}(xbc). \tag{17}$$

Differentiating the right side with regard to b_i gives

$$s(s-1)\Delta^{s-2}B_i(xbc) + s\Delta^{s-1}(cx)_{jk},$$

whence, after multiplying by an arbitrary y_i and summing for $i=1, 2, 3$,

$$\left(y\,\middle|\,\frac{\partial}{\partial b}\right)\left(x\,\middle|\,\frac{\partial}{\partial a}\right)\Delta^s = s(s-1)\Delta^{s-2}(ayc)(xbc) + s\Delta^{s-1}(xyc). \tag{18}$$

But $(ayc)(xbc) - (axc)(ybc) = (abc)(xyc) = \Delta(xyc)$ identically (p. 42, (29)). Hence after rewriting (18) with x, y interchanged throughout and subtracting from (18), we have

$$\begin{aligned}\left(x_y\,\middle|\,\frac{\partial}{\partial a}\;\frac{\partial}{\partial b}\right)\Delta^s &= s(s-1)\Delta^{s-2}\Delta(xyc) + 2s\Delta^{s-1}(xyc) \\ &= s(s+1)\Delta^{s-1}(xyc). \quad .\;.\;.\;. \end{aligned} \tag{19}$$

---

[1] Cf. Grace and Young, *The Algebra of Invariants* (Cambridge, 1903), p. 259.

Next, in the same way,

$$\left(z \,\middle|\, \frac{\partial}{\partial c}\right)\left(xy \,\middle|\, \frac{\partial}{\partial a}\ \frac{\partial}{\partial b}\right)\Delta^s$$

$$= s(s+1)\,(s-1)\,\Delta^{s-2}\,(abz)\,(xyc) + s(s+1)\,\Delta^{s-1}\,(xyz). \quad (20)$$

Interchanging x, y, z according to the scheme $\overline{xy}, \overline{z}$ which is

$$xy,\ z \qquad yz,\ x \qquad zx,\ y$$

we obtain three such equations as (20): and the result of adding them up is (p. 42, (28))

$$\left(xyz \,\middle|\, \frac{\partial}{\partial a}\ \frac{\partial}{\partial b}\ \frac{\partial}{\partial c}\right)\Delta^s = s\,(s+1)\,(s-1)\,\Delta^{s-2}\,(abc)\,(xyz)$$
$$+ 3s\,(s+1)\,\Delta^{s-1}\,(xyz)$$

since the last term is the same for each, $(xyz) = (yzx) = (zxy)$. Thus

$$(xyz)\,\Omega\Delta^s = s\,(s+1)\,(s-1+3)\,\Delta^{s-1}\,(xyz)$$

or

$$\Omega\Delta^s = s\,(s+1)\,(s+2)\,\Delta^{s-1}. \quad . \quad . \quad . \quad (21)$$

EXAMPLES

1. Prove by this method that if $\Delta = (ab\ldots ef\ldots m)$ is a determinant of order n, then

$$\left(x \,\middle|\, \frac{\partial}{\partial a}\right)\Delta^s = s\Delta^{s-1}\,(xbc\ldots m), \quad \left(xy \,\middle|\, \frac{\partial}{\partial a}\ \frac{\partial}{\partial b}\right)\Delta^s = s(s+1)\,\Delta^{s-1}(xyc\ldots m).$$

2. For $k < n$, if $x\ldots y$ and $a\ldots e$ both denote k columns, prove

$$\left(x\ldots y \,\middle|\, \frac{\partial}{\partial a}\cdots\frac{\partial}{\partial e}\right)\Delta^s = s\,(s+1)\ldots(s+k-1)\,\Delta^{s-1}\,(x\ldots y\,f\ldots m).$$

[Use induction, and proceed as before.]

3. If $k = n$, deduce the theorem

$$\Omega\Delta^s = s(s+1)\ldots(s+n-1)\,\Delta^{s-1}.$$

4. If $r < s$, prove

$$\Omega^r\Delta^s = \frac{(s+n-1)!}{(s-1)!}\,\frac{(s+n-2)!}{(s-2)!}\cdots\frac{(s+n-r)!}{(s-r)!}\,\Delta^{s-r}.$$

5. If $r = s$, this becomes

$$\Omega^s\Delta^s = n!\,\frac{(n+1)!}{1!}\,\frac{(n+2)!}{2!}\cdots\frac{(n+s-1)!}{(s-1)!}.$$

6. If $r > s$, prove $\Omega^r\Delta^s = 0$.

7. Prove $\left| \frac{\partial}{\partial a_i} \frac{\partial}{\partial b_j} \cdot \frac{\partial}{\partial e_k} \quad a_1 b_2 \ldots m_n \right|^s$

$$= s(s+1)\ldots(s+k-1)\Delta^{s-1}\,|f_l \ldots m_n|,$$

where $|a_i b_j \ldots e_k|$ and $|f_l \ldots m_n|$ are complementary co-factors in Δ.

4. Theorem of Corresponding Matrices adapted to the Capelli Operator.

If x_a denote the operator $(x\,|\,\partial/\partial a)$, then the theorem of corresponding matrices yields

$$\begin{vmatrix} \left(x\middle|\frac{\partial}{\partial a}\right), & \left(y\middle|\frac{\partial}{\partial a}\right) \\ \left(x\middle|\frac{\partial}{\partial b}\right), & \left(y\middle|\frac{\partial}{\partial b}\right) \end{vmatrix} \equiv \begin{vmatrix} x_a & y_a \\ x_b & y_b \end{vmatrix} = \left(xy\middle|\frac{\partial}{\partial a}\,\frac{\partial}{\partial b}\right). \qquad (22)$$

This breaks down if the variables a, b are replaced by x, y or by functions of x and y, because the left-hand factor of each term of the expanded operator, $x_a y_b - x_b y_a$, acts on the right. Thus we find, if the determinant is expanded by columns,

$$\begin{matrix} y_x \\ y_y \end{matrix}\; f = \Sigma\,(xy)_{ij}\left(\frac{\partial}{\partial x}\,\frac{\partial}{\partial y}\right)_{ij} f - x_x f. \qquad (23)$$

The second term on the right, $-x_x f$, is due to $\frac{\partial}{\partial y_i}$ in x_y acting directly on y_i in y_x. Also, if we expand by rows, writing the determinant as $x_x y_y - y_x x_y$ we obtain still another result. Let us therefore agree to *expand by columns* in each case when this ambiguity may arise.

Further, let the two letters of an element x_y be called the upper x and the lower y, so that the lower letter represents a set of differential operators $\frac{\partial}{\partial y_i}$. Then we notice that when a lower letter y in an earlier column is followed by the same y as upper letter in a later column, a new term, as already remarked, may arise. It is well to have names to distinguish these terms. In an operator such as the above, terms arising from direct differentiation upon f are called *extrinsic*, but terms arising within the operator are *intrinsic*. In (23) above, $-x_x f$ is an intrinsic term, while the summation Σ gives a series of extrinsic terms.

By rearranging the terms of (23) we have the relation

$$\begin{matrix} y_x \\ y_y+1 \end{matrix} \quad f = \Sigma\,(xy)_{ij}\left(\frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\right)_{ij} f = \left(xy\quad \frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\right) f. \quad (24)$$

This ingenious device absorbs the intrinsic term into the operator by adding a new extrinsic term $x_x f$ through increase of the lower right element y_y by unity. It was Capelli[1] who first discovered this law of adjustment in its generality, which can take the elegant form for $r \leqslant n$ sets of variables $x, y, \ldots, z, t$,

$$\Delta = \left. \begin{matrix} x_x & y_x & \ldots & z_x & t_x \\ x_y & y_y+1 & \ldots & z_y & t_y \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ x_z & y_z & \ldots & z_z+r-2 & t_z \\ x_t & y_t & \ldots & z_t & t_t+r-1 \end{matrix} \right|$$

$$= \left(xy \ldots zt \,\middle|\, \frac{\partial}{\partial x}\ \frac{\partial}{\partial y} \cdots \frac{\partial}{\partial z}\ \frac{\partial}{\partial t}\right) = H. \quad (25)$$

Here the leading diagonal has 0, 1, 2, . . . , $r-1$ added to its respective elements, which otherwise agree with the algebraic theorem of corresponding matrices.

Proof.—If we expand Δ by a Laplace $(2 : r-2)$ development, every minor from the first two columns is of the required type, since those involving row$_2$ obey the law shown in (24), and all others have no intrinsic terms.

Accordingly we assume the theorem true for all minors of the first $r-1$ columns, and proceed to prove it inductively for Δ itself by the $(r-1 : 1)$ development. On performing this expansion of Δ, we have

$$\Delta = T_1 t_x + T_2 t_y + \ldots + T_{r-1} t_z + T_r (t_t + r - 1), \quad (26)$$

where T_i is the co-factor of the last element in row$_i$. But by hypothesis

$$T_1 t_x = (-)^{r-1}\left(xy \ldots z \,\middle|\, \frac{\partial}{\partial y} \cdots \frac{\partial}{\partial z}\ \frac{\partial}{\partial t}\right) t_x.$$

Here the only intrinsic terms are due to the presence of t in t_x.

---

[1] *Math. Annalen*, **29** (1887), 331–338

But $\frac{\partial}{\partial t_i} t_x = \frac{\partial}{\partial x_i}$. Hence, by summing $i = 1, 2, \ldots, n$, the intrinsic terms of $T_1 t_x$ combine into the single expression

$$(-)^{r-1}\left(xy \ldots z \,\middle|\, \frac{\partial}{\partial y} \ldots \frac{\partial}{\partial z}\, \frac{\partial}{\partial x}\right)$$

$$= -\left(xy \ldots z \;\; \frac{\partial}{\partial x}\, \frac{\partial}{\partial y} \ldots \frac{\partial}{\partial z}\right) = -T_r$$

on shifting $\frac{\partial}{\partial x}$ through $r-2$ places. Similarly each of the co-factors $T_2, \ldots, T_{r-1}$ furnish $-T_r$ as intrinsic term. So the sum of all intrinsic terms in Δ cancels the $T_r(r-1)$ in the last term of (26) which itself is free from intrinsic terms. Hence we can write

$$\Delta = T_1{}' t_x + \ldots + T_r{}' t_t,$$

where the accent denotes that the operation passes over $t_x, \ldots, t_t$ and acts only on what may follow. Collecting terms we now have

$$\Delta = \left(xy \ldots zt \,\middle|\, \frac{\partial}{\partial x}\, \frac{\partial}{\partial y} \ldots \frac{\partial}{\partial z}\, \frac{\partial}{\partial t}\right),$$

which proves the theorem.

Corollary I.—*If* r > n, Δ *vanishes identically.*

Corollary II.—Δ *is unchanged by deranging* x, y, . . . , z, t *similarly in* both *rows and columns. For this leaves* H *unchanged.*

Corollary III.—Δ *is unchanged by transposition, followed by reversal of the integers to* r − 1, r − 2, . . . , 3, 2, 1, 0 *in the leading diagonal.*

This follows by induction proceeding from the final column towards the left.

EXAMPLES

1. Prove $\begin{vmatrix} x_x + 2 & x_y & x_z \\ y_x & y_y + 1 & y_z \\ z_x & z_y & \end{vmatrix} = \Sigma (xyz)_{ijk} \left(\frac{\partial}{\partial x}\, \frac{\partial}{\partial y}\, \frac{\partial}{\partial z}\right)_{ijk}.$

2. Prove, if $n > 3$, and η is independent of x and y,

$$\begin{vmatrix} x_x + 2 & x_y & x_z \\ \eta_x & \eta_y & \eta_z \\ z_x & z_y & z_z \end{vmatrix} \qquad \left(x\eta z \,\middle|\, \frac{\partial}{\partial x}\, \frac{\partial}{\partial y}\, \frac{\partial}{\partial z}\right),$$

3. The general Capelli operator $\left(\xi y \zeta t \omega \,\middle|\, \frac{\partial}{\partial x}\,\frac{\partial}{\partial y}\,\frac{\partial}{\partial z}\,\frac{\partial}{\partial t}\,\frac{\partial}{\partial u}\right)$ involving some upper and lower letters y, t alike, and others entirely independent, is equal to a determinant of type (25) with 0, 1, 0, 3, 0 replacing 0, 1, 2, . . . , in the leading diagonal terms.

4. If p, q satisfy all the laws of §1, p. 57, except that of commutative multiplication which is replaced by

$$pq - qp = 1,$$

prove that $p^2q^2 = pq(pq+1)$, $p^3q^3 = pq(pq+1)(pq+2)$,

$$q^2p^2 = qp(qp-1),\quad q^3p^3 = qp(qp-1)(qp-2).$$

[Try first when $p = \frac{d}{dq}$. Next try directly by substituting $pq - 1$ for qp in $p(qp)q$].

5. Prove
$$p^rq^r = pq(pq+1)\ldots(pq+r-1)$$
$$q^rp^r = qp(qp-1)\ldots(qp-r+1).$$

6. Prove $p^2q - qp^2 = 2p$, $p^3q - qp^3 = 3p^2$, $p^nq - qp^n = np^{n-1}$.

5. Connexion between Substitutional Analysis and Differentiation.

The preceding investigations show that a close analogy exists between the typical process of algebra, the *permutation*, and that of analysis, *differentiation*. Indeed many of the properties of matrices, determinants, and the like, are rendered the clearer by bringing into play this twofold aspect of what is really one fundamental operation. A very simple example will suffice to lead up to the general idea.[1] Consider the operation of $\frac{d}{dx}$ upon x^n when n is a positive integer. If we write x^n as a product of n factors each equal to x,

$$xxxx\ldots,$$

it is clear that we can pick out an x in n different ways: we can then substitute unity for this factor in n different ways. If we do so, and add up the results we arrive at nx^{n-1}, namely the result of operating with $\frac{d}{dx}$ on x^n. Thus

$$\frac{d}{dx}x^n = 1xxx\ldots + x1xx\ldots + \ldots + xxx\ldots 1 = nx^{n-1}. \tag{27}$$

Similarly

$$y\frac{d}{dx}x^n = yxxx\ldots + xyxx\ldots + \ldots + xxx\ldots y = nyx^{n-1}. \tag{28}$$

---

[1] Cf. Macmahon, *Combinatory Analysis*, I (Cambridge, 1915), p. 224.

Here the left-hand operator is the simplest type of polar operator; and we see from the series to which it gives rise that it is essentially an operation involving permutations of substitutions.

Now the determinantal permutation

$$\dot{a}\dot{b},\ \dot{c} \equiv ab,\ c \qquad bc,\ a \qquad ca,\ b$$

which takes its rise in the Laplace development of a determinant,

$$\Delta = |\,a_1 b_2 c_3\,| = |\,\dot{a}_1 \dot{b}_2\,|\,\dot{c}_3 = |\,a_1 b_2\,|\,c_3 + |\,b_1 c_2\,|\,a_3 + |\,c_1 a_2\,|\,b_3,$$

has the same general features, only complicated by the change of sign which accompanies an interchange of letters. And if the determinantal permutation operates on letters representing columns of determinants, it is found in all cases to be expressible by differential operators. For example, the process $\dot{a}\dot{b},\ \dot{c}$ applied to a product of determinants of any, the same, order, say the fourth order,

$$(\dot{a}\dot{b}de)\ (\dot{c}fgh)$$

may be equated to

$$\tfrac{1}{2}\left(abc\ \ \frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\ \frac{\partial}{\partial z}\right)(xyde)\ (zfgh). \tag{29}$$

For this operator is equal to

$$\tfrac{1}{2}\left(ab\ \ \frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\right)\left(c\,\middle|\,\frac{\partial}{\partial z}\right) + \tfrac{1}{2}\left(bc\ \ \frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\qquad \frac{\partial}{\partial z}\right)$$
$$+\ \tfrac{1}{2}\left(ca\,\middle|\,\frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\right)\left(b\ \ \frac{\partial}{\partial z}\right). \tag{30}$$

Also by (14),

$$ab\left|\frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\right)(xyde) = 2(abde) \qquad\qquad \frac{\partial}{\partial z}\Big)\,(zfgh) = (cfgh).$$

Hence the effect of the whole operation $\left(abc\,\middle|\,\frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\ \frac{\partial}{\partial z}\right)$ is

$$2(abde)\ (cfgh) + 2(bcde)\ (afgh) + 2(cade)\ (bfgh),$$

which gives the required result.

In general, the permutation operations of §11, p. 47 which lead to the fundamental identities can be expressed as differential operators. For example, the series of $\binom{i+j}{i}$ terms, given by

$(\dot{A}_i L_{n-i})(\dot{B}_j M_{n-j})$ can be generated from a single product of two n-rowed determinants

$$(x_1 x_2 \ldots x_i L_{n-i})\ (y_1 y_2 \ldots y\ M_{n-j})$$

by the operator

$$\frac{1}{i!\,j!}\left(A_i B_j\ \frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_i}\ \frac{\partial}{\partial y_1} \cdots \frac{\partial}{\partial y}\right),$$

where each x denotes a column of the determinant, A_i denotes i columns $a_1, a_2, \ldots, a_i$, and B_j, j columns, while

$$\left(A_i B_j \middle| \frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial y_j}\right) = \sum_K (a_1 \ldots a_i b_1 \ldots b\ )_K \left(\frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial y_j}\right)_K,$$

K being a row of $i+j$ different suffixes chosen in $\binom{n}{i+j}$ different ways from the integers 1, 2, ..., n.

Just as A_i is short for the $i \times n$ matrix $a_1 a_2 \ldots a_i$ let $\dfrac{\partial}{\partial X_i}$ be short for $\dfrac{\partial}{\partial x_1} \cdots \dfrac{\partial}{\partial x_i}$, and $\dfrac{\partial}{\partial Y_j}$ for $\dfrac{\partial}{\partial y_1} \cdots \dfrac{\partial}{\partial y_j}$. Then we have the relation

$$(\dot{A}_i L)(\dot{B}_j M) = \frac{1}{i!\,j!}\left(A_i B_j\ \frac{\partial}{\partial X_i}\ \frac{\partial}{\partial Y_j}\right)(X_i L)(Y_j M). \qquad (31)$$

The proof follows the lines of the previous special case. For as in (29), (30)

$$\left(A_i B_j \middle| \frac{\partial}{\partial X_i}\ \frac{\partial}{\partial Y_j}\right) = \left(\dot{A}_i \middle| \frac{\partial}{\partial X_i}\right)\left(\dot{B}_j\ \frac{\partial}{\partial Y_j}\right),$$

and by (16)

$$\left(A_i \middle| \frac{\partial}{\partial X_i}\right)(X_i L) = i!\,(A_i L).$$

Similarly for B, Y. This at once yields the result.

Incidentally this affords an alternative proof of Sylvester's theorem (§9, (II), p. 45) when $i + j = n$, because the matrix product operator then factorizes as

$$(A_i B_j)\left(\frac{\partial}{\partial X_i}\ \frac{\partial}{\partial Y_j}\right),$$

showing that $(A_i B_j)$ is a factor of the series $(\dot{A}_i L)(\dot{B}\ M)$.

Once more, by the same reasoning, for several matrices A_i, B_j, C_k . . ., if $i + j + k + \ldots \leqslant n$, we may write

$$(\dot{A}L)(\dot{B}M)(\dot{C}N)\ldots$$
$$= \frac{1}{i!\, j! \ldots}\left(ABC \ldots \quad \frac{\partial}{\partial X}\, \frac{\partial}{\partial Y}\, \frac{\partial}{\partial Z} \ldots\right)(XL)\,(YM)\,(ZN)\ldots$$

where the currencies of

$$\begin{array}{ccc} A & X & L \\ B & Y & M \\ C & Z & N \end{array}$$

are

$$\begin{array}{ccc} i & i & n - i \\ j & j & n - j \\ k & k & n - k \end{array}$$

respectively.

An important particular case of the above is the following type of identity, involving the Cayley operator and a product of factors $a_x, b_y, c_z \ldots$ where $a_x = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n$.

$$\left.\begin{array}{lll} \text{If} & n = 2, & \Omega a_x b_y = (ab); \\ \text{if} & n = 3, & \Omega a_x b_y c_z = (abc); \\ \text{if} & n = 4, & \Omega a_x b_y c_z d_t = (abcd), \text{ \&c.} \end{array}\right\} \quad . \quad (32)$$

Thus, if $n = 3$, $\Omega = \Sigma \pm \dfrac{\partial}{\partial x_1}\, \dfrac{\partial}{\partial y_2}\, \dfrac{\partial}{\partial z_3}$, whence $\Omega\, a_x\, b_y\, c_z = \Sigma \pm a_1 b_2 c_3 = (abc)$. And if there were more than three such factors, the result would contain several terms, with a determinant like (abc) appearing in each. For instance, still with $n = 3$,

$$\Omega\, a_x b_y c_z d_x = \Sigma \pm \frac{\partial}{\partial x_1}\, \frac{\partial}{\partial y_2}\, \frac{\partial}{\partial z_3}\, a_x b_y c_z d_x$$
$$= (abc)\, d_x + (dbc)\, a_x. \quad . \quad . \qquad (33)$$

Evidently the process mimics the ordinary rule of differentiating a product (cf. (27) and (28)).

EXAMPLES

1. If $\Omega = \Sigma \pm \frac{\partial}{\partial x_1} \frac{\partial}{\partial y_2} \frac{\partial}{\partial z_3}$, prove

$$\Omega a_x b_y c_z a'_x b'_y = (abc) a'_x b'_y + (a'bc) a_x b'_y + (ab'c) a'_x b_y + (a'b'c) a_x b_y.$$

2. $\Omega a_x^2 b_y^2 c_z = 4(abc) a_x b_y$.

3. $\Omega a_x^m b_y^n c_z^p = mnp(abc) a_x^{m-1} b_y^{n-1} c_z^{p-1}$.

4. $\Omega^r a_x^m b_y^n c_z^p = \frac{m!\, n!\, p!}{(m-r)!\,(n-r)!\,(p-r)!} (abc)^r a_x^{m-r} b_y^{n-r} c_z^{p-r}$.

5. Give the corresponding identity for a determinant Ω of the nth order.

6. If the n^2 elements of $\Delta = (abc \ldots m)$ are functions of a single variable t, and an accent denotes differentiation with regard to t of elements in the column indicated, prove

$$\frac{d\Delta}{dt} = (a'bc \ldots m) + (ab'c \ldots m) + (abc' \ldots m) + \ldots + (abc \ldots m').$$

7. If $S = a_0 + a_1 \Delta + a_2 \Delta^2 + \ldots + a_p \Delta^p + \ldots$ and $\Delta = |x_1 y_2 \ldots t_n|$ where the coefficients a_i are constants and the series is convergent, prove

$$\Omega S = n! \left\{ a_1 + (n+1) a_2 \Delta + \frac{(n+1)(n+2)}{1 \, . \, 2} a_3 \Delta^2 + \ldots \right\}.$$

8. Prove (i) $\Omega \frac{1}{1-\Delta} = n! \, \frac{1}{(1-\Delta)^{n+1}}$.

(ii) $\Omega \log(1-\Delta) = -(n-1)! \, [(1-\Delta)^{-n} - 1]/\Delta$.

9. $\left(\frac{}{\partial x_1 \partial y_2} - \frac{\partial^2}{\partial x_2 \partial y_1} \right) \frac{}{x_1 y_2 - x_2 y_1} = 0.$

10. $\left(\frac{\partial^2}{\partial x_1 \partial y_2} - \frac{}{\partial x_2 \partial y_1} \right) (x_1 y_2 - x_2 y_1)^{-s}$

$$= -s(-s+1)(x_1 y_2 - x_2 y_1)^{-s-1}.$$

11. Prove $\Omega \frac{1}{\Delta} = 0$ for all orders of the determinant.

12. Prove the Cayley formula (§3, p. 114) true for negative values of s.

13. The partial differential equation

$$\frac{\partial^2 z}{\partial x_1 \partial y_2} \qquad \frac{\partial^2 z}{\partial x_2 \partial y_1}$$

is satisfied by

$$z = c \left\{ 1 + \frac{\Delta}{2!} + \frac{\Delta^2}{2!\,3!} + \frac{\Delta^3}{3!\,4!} + \ldots \right\}$$

where $\Delta = x_1 y_2 - x_2 y_1$.

14. The partial differential equation in nine independent variables $x_1, \ldots, z_3$

$$\Sigma \pm \frac{\partial^3 V}{\partial x_1 \partial y_2 \partial y_3} = V, \quad \text{or briefly} \quad \Omega V = V$$

is satisfied by

$$V = c\left\{1 + \frac{\Delta}{3!} + \frac{\Delta^2}{3!\,4!} + \frac{2!\,\Delta^3}{3!\,4!\,5!} + \dots\right\}$$

For n^2 independent variables,

$$V = c\left\{1 + \frac{\Delta}{n!} + \frac{\Delta^2}{n!(n+1)!} + \frac{2!\,\Delta^3}{n!(n+1)!(n+2)!} + \frac{2!\,3!\,\Delta^4}{n!(n+1)!(n+2)!(n+3)!} + \dots\right\}.$$

15. Prove $\left(a\,\middle|\,\frac{\partial}{\partial x}\right)\log\Delta = \frac{(ayz\dots t)}{(xyz\dots t)}$, $\left(ab\,\middle|\,\frac{\partial}{\partial x}\,\frac{\partial}{\partial y}\right)\log\Delta = \frac{(abz\dots t)}{(xyz\dots t)}$, $\left(abc\,\middle|\,\frac{\partial}{\partial x}\,\frac{\partial}{\partial y}\,\frac{\partial}{\partial z}\right)\log\Delta = 2!\,\frac{(abc\dots t)}{(xyz\dots t)}$, and finally

$$\Omega\log\Delta = \frac{(n-1)!}{\Delta},\quad \Omega^2\log\Delta = 0.$$

16. The solution of n linear equations for $\xi, \eta, \dots, \omega$

$$x_i\xi + y_i\eta + \dots + t_i\omega = a_i,\quad i = 1, 2, \dots, n$$

is given by

$$\xi = \left(a\,\middle|\,\frac{\partial}{\partial x}\right)\log\Delta,\quad \eta = \left(a\,\middle|\,\frac{\partial}{\partial y}\right)\log\Delta,\ \&c.$$

6. Jacobians.

The determinant $|\partial u_i/\partial x_j|$ of n rows (i) and columns (j), whose elements are the n^2 first partial differential coefficients of n functions $u_1, u_2, \dots, u_n$ with regard to their n independent arguments $x_1, x_2, \dots, x_n$, is called the Jacobian of the set u with regard to the set x. It can be denoted in various ways:

$$\frac{\partial u_i}{\partial x_j} = \frac{\partial(u_1, u_2, \dots, u_n)}{\partial(x_1, x_2, \dots, x_n)} = \frac{\partial(u)}{\partial(x)} = \Delta.$$

Its properties are essentially algebraic, once the fundamental facts of partial differentiation are assumed, and in particular the theorems: if ϕ is an explicit function of r arguments $u_1\ u_2, \dots, u_r$, then

$$\frac{\partial\phi}{\partial x_i} = \frac{\partial\phi}{\partial u_1}\,\frac{\partial u_1}{\partial x_i} + \frac{\partial\phi}{\partial u_2}\,\frac{\partial u_2}{\partial x_i} + \dots + \frac{\partial\phi}{\partial u_r}\,\frac{\partial u_r}{\partial x_i};$$

and if $\psi(u_1, u_2, \dots, u_r, x_1, x_2, \dots, x_n) = 0$, then

$$\frac{\partial\psi}{\partial u_1}\,\frac{\partial u_1}{\partial x_i} + \dots + \frac{\partial\psi}{\partial u_r}\,\frac{\partial u_r}{\partial x} + \frac{\partial\psi}{\partial x} = 0.$$

The chief properties of the Jacobian are contained in the following six theorems.

I. *If the* u's *are explicit functions of* $y_1, y_2, \ldots, y_n$ *which in turn are explicit functions of the* x's, *then*

$$\frac{\partial(u)}{\partial(x)} = \frac{\partial(u)}{\partial(y)}\,\frac{\partial(y)}{\partial(x)}.$$

For by multiplication the (i, j)th element in the product determinant is

$$\frac{\partial u_i}{\partial y_1}\,\frac{\partial y_1}{\partial x_j} + \ldots + \frac{\partial u_i}{\partial y_n}\,\frac{\partial y_n}{\partial x_j} = \frac{\partial u_i}{\partial x_j}.$$

II. *If the* n *equations* $u_i = u_i(x_1, \ldots, x_n)$ *can be solved for the* x's *in terms of the* u's, *and the Jacobian* $|\partial x_i/\partial u_j|$ *can be constructed, then*

$$\frac{\partial(x)}{\partial(u)}\,\frac{\partial(u)}{\partial(x)} = 1.$$

For the (i, j)th element in the product is $\dfrac{\partial x_i}{\partial x_j} = \delta_{ij}$, leading to the unit determinant (§2, p. 32).

III. *If* $F_i(u_1, u_2, \ldots, u_n, x_1, x_2, \ldots, x_n) = 0;\ i = 1, 2, \ldots, n,$ *then*

$$\frac{\partial(u)}{\partial(x)} = (-)^n \frac{\partial(F)}{\partial(x)}\Big/\frac{\partial(F)}{\partial(u)}.$$

For by actual multiplication

$$\begin{aligned}\frac{\partial(u)}{\partial(x)}\,\frac{\partial(F)}{\partial(u)} = \left|\frac{\partial F_i}{\partial u_1}\,\frac{\partial u_1}{\partial x_j} + \ldots + \frac{\partial F_i}{\partial u_n}\,\frac{\partial u_n}{\partial x_j}\right| \\ = \left|-\frac{\partial F_i}{\partial x}\right| = (-)^n \frac{\partial(F)}{\partial(x)}.\end{aligned}$$

IV. **Jacobi's Lemma.**[1]—*If* $A_1, A_2, \ldots, A_n$ *are the co-factors of* $\dfrac{\partial u_i}{\partial x_1}, \dfrac{\partial u_i}{\partial x_2}, \ldots, \dfrac{\partial u_i}{\partial x_n}$ *in the Jacobian, then*

$$\frac{\partial A_1}{\partial x_1} + \frac{\partial A_2}{\partial x_2} + \ldots + \frac{\partial A_n}{\partial x_n} = 0$$

identically.

---

[1] Crelle, **27**, (1844), 201–209. *Collected Works*, **4**, 317.

For $\partial A_1/\partial x_1$ consists of a sum of $n-1$ determinants, due to differentiating in turn the $n-1$ columns of the co-factor A_1. We can then arrange all the $n(n-1)$ terms arising from the n differential coefficients $\partial A_i/\partial x_i$ as a *skew* symmetric matrix of order n, with terms arising from A_i arranged in the ith row. Since the matrix is skew, the sum of its terms vanishes. It is left as an exercise for the reader to develop this proof.

7. Rank of Jacobian Matrix.

The square matrix $[\partial u_i/\partial x_j]$ of order n is the Jacobian matrix, and its rank r is the highest order of a minor determinant Δ_r, which does not vanish identically. What follows now is closely analogous to the theorem on p. 73.

V. *If a functional relation* $\phi(\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_\mathrm{p}) = 0$ *connects* p *of the* u*'s, then every minor* $\partial(\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_\mathrm{p})/\partial(\mathrm{x}_\mathrm{i}, \ldots, \mathrm{x}_\mathrm{j})$ *of order* p *involving these* u*'s vanishes identically.*

For since $0 = \dfrac{\partial \phi}{\partial x_i} = \dfrac{\partial \phi}{\partial u_1}\,\dfrac{\partial u_1}{\partial x_i} + \ldots + \dfrac{\partial \phi}{\partial u_p}\,\dfrac{\partial u_p}{\partial x_i}$, we have p linear equations from which $\dfrac{\partial \phi}{\partial u_1}, \ldots, \dfrac{\partial \phi}{\partial u_p}$ may be eliminated, so giving the desired result. Incidentally the rank r is necessarily less than n.

VI. *If* $\Delta_\mathrm{r} \equiv \dfrac{\partial(\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_\mathrm{r})}{\partial(\mathrm{x}_1, \mathrm{x}_2, \ldots, \mathrm{x}_\mathrm{r})} \neq 0$, *where* r *is the rank of the Jacobian matrix, then the functions* $\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_\mathrm{r}$ *are independent, while each of the* $\mathrm{n}-\mathrm{r}$ *remaining* u*'s is expressible in terms of* $\mathrm{u}_1, \mathrm{u}_2, \ldots, \mathrm{u}_\mathrm{r}$.

By V we already know that $u_1, \ldots, u_r$ are independent, otherwise Δ_r vanishes. So we take these r together with $x_{r+1}, x_{r+2}, \ldots, x_n$ as new independent variables and express the remaining x's as

$$x = \phi_j(u_1, \ldots, u_r, x_{r+1}, \ldots, x_n), \quad j = 1, 2, \ldots, r.$$

Also let $\quad u_i = f_i(x_1, x_2, \ldots, x_n), \quad i = 1, 2, \ldots, n.$

Then by differentiation of $u_1, u_2, \ldots, u_n$ (as functions of the new independent variables) with regard to $x_s\,(s > r)$,

$$0=\frac{\partial f_1}{\partial x_1}\,\frac{\partial \phi_1}{\partial x_s}+\dots+\frac{\partial f_1}{\partial x_r}\,\frac{\partial \phi_r}{\partial x_s}+\frac{\partial f_1}{\partial x_s},$$

$$\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot$$

$$0=\frac{\partial f_r}{\partial x_1}\,\frac{\partial \phi_1}{\partial x_s}+\dots+\frac{\partial f_r}{\partial x_r}\,\frac{\partial \phi_r}{\partial x_s}+\frac{\partial f_r}{\partial x_s},$$

$$\frac{\partial u_k}{\partial x_s}=\frac{\partial f_k}{\partial x_1}\,\frac{\partial \phi_1}{\partial x_s}+\dots+\frac{\partial f_k}{\partial x_r}\,\frac{\partial \phi_r}{\partial x_s}+\frac{\partial f_k}{\partial x_s},$$

$$k=r+1,\quad r+2,\dots,n.$$

Eliminating $\partial\phi_1/\partial x_s,\dots,\partial\phi_r/\partial x_s$, we obtain

$$\begin{array}{llll}
\dfrac{\partial f_1}{\partial x_1}, & & \dfrac{\partial f_1}{\partial x_r}, & \dfrac{\partial f_1}{\partial x_s} \\
\cdot\quad\cdot\quad\cdot & \cdot\quad\cdot\quad\cdot & \cdot\quad\cdot & \cdot \\
\dfrac{\partial f_r}{\partial x_1}, & \dots, & \dfrac{\partial f_r}{\partial x_r}, & \dfrac{\partial f_r}{\partial x_s} \\
\dfrac{\partial f_k}{\partial x_1}, & \dots, & \dfrac{\partial f_k}{\partial x_r}, & \dfrac{\partial f_k}{\partial x_s}-\dfrac{\partial u_k}{\partial x_s}
\end{array}=0.$$

Expanding by the last row,

$$\frac{\partial(f_1,\dots,f_r,f_k)}{\partial(x_1,\dots,x_r,x_s)}=\frac{\partial u_k}{\partial x_s}\,\frac{\partial(f_1,\dots,f_r)}{\partial(x_1,\dots,x_r)}.$$

But by hypothesis the left-hand member of this identity vanishes, whereas $\partial(f_1,\dots,f_r)/\partial(x_1,\dots,x_r)$ does not. Hence $\partial u_k/\partial x_s$ vanishes identically, so that u_k is independent of x_s. But s can be any of $(r+1),\dots,n$. Hence u_k is expressible entirely as a function of the remaining variables $u_1, u_2, \dots, u_r$: and this proves the theorem.

Corollary.—*Provided its rank is* r *the matrix* $[\partial u_i/\partial x_j]$ *need not necessarily be square.*

In fact the above proof holds when the number of u's is m, where $m \geqslant r$.

CHAPTER VIII

BINARY FORMS

1. Binary Invariants.

We shall now consider, as a preliminary to more complicated structures, a particular type of expression called the binary form: and this will be dealt with broadly in the order suggested by the history of algebra since the time when Lagrange and Gauss hinted at properties of linear transformations, finally to be disclosed in an epoch-making publication by Boole in November, 1841.

Let us consider this partly from a geometrical point of view. Suppose that

$$F(x, y) \equiv ax^2 + 2hxy + by^2 + 2gx + 2fy + c, \quad . \quad (1)$$

equated to zero, represents a conic referred to Cartesian co-ordinates (x, y), and that a change of axes is made, as indicated by the equations,

$$T: \quad \left.\begin{aligned} x &= l_1 x' + m_1 y' \\ y &= l_2 x' + m_2 y' \end{aligned}\right\} \quad . \quad . \quad . \qquad (2)$$

for the old co-ordinates x, y in terms of the new, x', y'. The origin remains fixed, and the only condition imposed on l_1, m_1, l_2, m_2 is the inequality

$$|M| \equiv l_1 m_2 - l_2 m_1 \neq 0. \quad . \quad . \quad . \quad . \quad (3)$$

This is a *homogeneous linear transformation* from x, y to x', y'. Frobenius and other writers, with no geometrical purpose immediately in view, call it a *substitution* rather than a transformation.

First we remark that any function of x, y may be expressed as a function of x', y'. Let us write

$$F(x, y) = F'(x', y') \qquad (4)$$

to denote the two aspects of this function. Next if we collect terms of F in groups U_r homogeneous in x, y, as indicated by the suffix r, we have in this case

$$U_2 + U_1 + U_0 = U_2' + U_1' + U_0'. \quad . \quad . \quad . \quad (5)$$

and in particular

$$U_2 = U_2', \quad U_1 = U_1', \quad U_0 = U_0'.$$

Thus $U_2 = ax^2 + 2hxy + by^2$

$$= a(l_1 x' + m_1 y')^2 + 2h(l_1 x' + m_1 y')(l_2 x' + m_2 y') + b(l_2 x' + m_2 y')^2$$
$$= a' x'^2 + 2h' x' y' + b' y'^2 = U_2',$$

provided

$$\begin{aligned} a' &= a l_1^2 + 2h l_1 l_2 + b l_2^2, \\ h' &= a l_1 m_1 + h(l_1 m_2 + l_2 m_1) + b l_2 m_2, \quad . \quad . \quad (6) \\ b' &= a m_1^2 + 2h m_1 m_2 + b m_2^2. \end{aligned}$$

It is at this point that the far-reaching result disclosed by Boole may be seen. Boole remarked that the discriminant $a'b' - h'^2$ of the quadratic U_2' reproduced that of U_2 together with a factor depending only on the coefficients of the transformation T. Namely,

$$a'b' - h'^2 = (l_1 m_2 - l_2 m_1)^2 (ab - h^2) \quad . \quad . \quad (7)$$

as is quite easy to verify. Let this be written

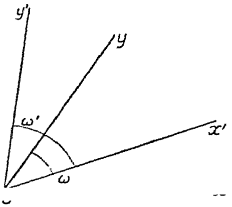
$$I(a') = | M |^2 . I(a), \quad . \quad . \quad . \quad . \quad . \quad . \quad (8)$$

where the contracted functional notation is adopted for brevity.

The factor $| M |$ is called the *modulus* or *determinant of the transformation* T.

Since this function $ab - h^2$ as a whole emerges unchanged in structure, but for a factor $| M |^2$ independent of the three arguments a, h, b of the function, it is called an *invariant* of the transformation. More precisely this is called a *relative*, to distinguish it from an *absolute*, invariant, because cases occur in which $I(a)$ and $I(a')$ are absolutely equal without the help of an extraneous factor $| M |^2$.

The significance of result (7) is better seen if the previous conditions (6) are studied. Each new coefficient a', h', b' is a complicated linear function of the old coefficients a, h, b; and



only this particular expression $a'b'-h'^2$, or a product involving this, turns out to have $ab-h^2$ as a factor—a property which shall be proved later.

Boole also found other interesting results which may shortly be stated. If ω denote the angle between the axes Ox, Oy, and ω' that between Ox', Oy', then the transformation T may be regarded as a means of referring the same geometrical figure to two sets of axes Oxy, $Ox'y'$ at the same origin. The assumption already made that M differs from zero implies

$$\sin\omega \neq 0, \quad \sin\omega' \neq 0,$$

i.e.

$$\omega \neq 0 \bmod \pi, \quad \omega' \neq 0 \bmod \pi.$$

The axes Ox, Oy are inclined to one another; and so are Ox', Oy'. Boole found the following relations:

$$|M| = \frac{\sin\omega}{\sin\omega'}, \quad \frac{a'b'-h'^2}{\sin^2\omega'} = \frac{ab-h^2}{\sin^2\omega}. \qquad (9)$$

$$\frac{a'+b'-2h'\cos\omega'}{\sin^2\omega'} = \frac{a+b-2h\cos\omega}{\sin^2\omega}. \qquad (10)$$

Here again is an instance of invariants: this time the invariant functions involve a, b, h, ω, which is a more complicated set of arguments than a, b, h alone. On the other hand the invariants are absolute, not merely relative.

2. Orthogonal Transformation and Invariants.

If we impose the conditions

$$\left.\begin{aligned} l_1^2+m_1^2 = l_2^2+m_2^2 = 1 \\ l_1 l_2 + m_1 m_2 = 0 \end{aligned}\right\} \quad . \quad . \quad . \qquad (11)$$

upon the coefficients of the binary linear transformation T we call it now an *orthogonal transformation.* Geometrically these conditions show that if the lines Ox, Oy are at right angles, so also are Ox', Oy': and conversely.

It will be seen that the values of the angles ω, ω' are now

$$\omega = \pm\omega' = \pm\frac{\pi}{2} \qquad (12)$$

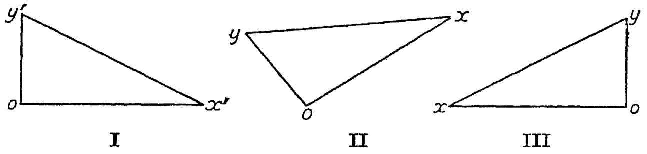
for if we write $l_1 = \cos\theta$, $m_1 = \sin\theta$, $l_2 = \cos\phi$, $m_2 = \sin\phi$ to satisfy the first condition, then

$$0 = l_1 l_2 + m_1 m_2 = \cos(\theta - \phi).$$

Hence θ and ϕ differ by an odd number of right angles. This is covered by two alternatives:

$$\left.\begin{array}{lll} \text{Either} & l_1 = m_2 = \cos\theta, & -l_2 = m_1 = \sin\theta \\ \text{or} & l_1 = -m_2 = \cos\theta, & l_2 = m_1 = \sin\theta \end{array}\right\}. \qquad (13)$$

In the first case the axes Oxy are obtained from $Ox'y'$ by a clockwise rotation through an angle θ: in the second they require, besides, turning over, to bring Ox into coincidence with Ox' and simultaneously Oy with Oy'.



This set of congruent right-angled triangles illustrates the point. The origins are separated for clearness. All such coplanar triangles fall into two classes according as whether, by a rigid displacement in their own plane, they can be superimposed or not. In the figure, I and II are in one class, III is in the other.

Algebraically the classes are included in one statement, easily verified, that

$$|M|^2 = \begin{matrix} & m_1 \\ l_2 & m_2 \end{matrix} = 1; \qquad (14)$$

and they are distinguished by extracting the square root. Thus

if | M | = 1, Oxy, Ox′y′ *belong to the same class*;

if | M | = − 1, Oxy, Ox′y′ *belong to different classes.*

We observe that for both classes of orthogonal transformation there are two absolute invariants derived from the results (9) and (10), namely,

$$\begin{aligned} a' + b' &= a + b, \\ a'b' - h'^2 &= ab - h^2. \end{aligned} \qquad . \quad . \quad . \qquad (15)$$

The importance of the above simple and well-known division into classes lies in the fact that it typifies the general case of orthogonal transformations for a set of n variables, not merely two. If we call the classes right and left handed, for distinction, a frame of right-handed axes can be brought to coincide with another right-handed frame by a rigid displacement in its own space; but an extra dimension of space is needed for a *rigid* displacement to bring it into coincidence with a left-handed frame. We shall return to this in Chapter IX.

It might be supposed that the absolute invariants are more important than the relative, but such is not the case, as the sequel will show. Speaking generally, the absolute invariant corresponds to a part of geometry which is merely a special case of something more general. The relative invariant is the important one.

All these expressions have a geometrical significance. For instance, $U_2 = 0$ represents two straight lines through the origin. If $ab - h^2 = 0$ the lines coincide, and if this is so, no mere change of axes will separate them; consequently $a'b' - h'^2 = 0$. But we can go a little further, supposing two real frames of axes, real co-ordinates, and real coefficients, so that $| M |^2 > 0$. Thus $ab - h^2$, $a'b' - h'^2$ are both positive, or both zero, or both negative. This is illustrated by the conic given by $F = 0$, the cases of ellipse, parabola, and hyperbola answering to this threefold classification.

Also if $a + b$ vanishes and the axes are at right angles then $U_2 = 0$ represents a pair of lines at right angles. A change to other rectangular axes at the same origin leaves the property unchanged. This explains why $a + b$ is an orthogonal invariant.

3. Development of the Invariant Theory.

The discovery made by Boole in 1841 was soon reinforced by an almost accidental observation by Eisenstein of an invariant belonging to a binary quartic. At once this attracted the attention of Cayley, Sylvester, and Salmon. Four years later Cayley put the subject in a more important light by asking two significant questions: (i) whether these ideas could be extended to binary forms U_3, U_4, . . . , U_n of all orders, and (ii) how far it was possible to discover *all* such invariantive functions?

To these ends he invented a device which he called *hyperdeterminants*, not unlike the device of denoting chemical substances and reactions by symbols and equations. He exhibited the properties and behaviour of hyperdeterminants, making a practical working tool of them. Out of this calculus modern algebra may be said to have sprung.[1]

In answer to Cayley's first question we have the systematic development of binary forms. The functions U_2, U_3, U_4, . . . , U_n already introduced are called the *binary quadratic*, *cubic*, *quartic* (or biquadratic), . . . , n-ic. We shall find it useful to call the rational integral function of *order* n in its arguments, a *polynomial.* A homogeneous polynomial is a *form* or *quantic.*

The order n is the highest degree in which the arguments occur in a term of the polynomial.

4. The Binary Form or Quantic.

We write the binary n-ic, U_n, as

$$U_n = a_0 x^n + n a_1 x^{n-1} y + \binom{n}{2} a_2 x^{n-2} y^2 + \ldots + a_n y^n, \qquad (16)$$

which Cayley shortened to

$$(a_0, a_1, \ldots, a_n \between x, y)^n. \quad . \quad . \quad . \quad (17)$$

The binomial coefficients do not make this any less general than the corresponding form

$$p_0 x^n + p_1 x^{n-1} y + \ldots + p_n y^n \quad . \quad . \quad (18)$$

which is sometimes used. They possess several clear advantages, especially when what are called polar forms are used, as we shall see later on.

We assume the theorem that the equation $U_n = 0$ has a root, and consequently, by repeating the argument, that U_n itself has n linear factors. Namely

$$U_n = p_0 (x - \alpha y)(x - \beta y) \ldots (x - \lambda y). \quad . \quad (19)$$

The set of n quantities

$$\alpha, \beta, \ldots, \lambda$$

are the roots of the n-ic $U = 0$.

---

[1] See an enthusiastic remark to the British Association (1869) by Sylvester, recorded in *Collected Works,* **2,** p. 656.

On multiplying out and comparing coefficients in (18) and (19) we obtain the relations

$$\begin{aligned} \Sigma\alpha &= \alpha+\beta+\ldots+\lambda = -p_1/p_0, \\ \Sigma\alpha\beta &= \alpha\beta+\alpha\gamma+\ldots = p_2/p_0, \\ \Sigma\alpha\beta\gamma &= \alpha\beta\gamma+\alpha\beta\delta+\ldots = -p_3/p_0, \qquad (20) \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha\beta\gamma\ldots\lambda &= (-)^n p_n/p_0. \end{aligned}$$

Here $\Sigma\alpha\beta\ldots\kappa$ with r factors has $\binom{n}{r}$ terms, found by all the combinations of r different letters chosen from $\alpha, \beta, \ldots, \lambda$. These are the elementary symmetric functions of the roots. They are called *symmetric* because any derangement of the letters $\alpha, \beta, \ldots, \lambda$ makes no difference to the functions.

Manifestly any product $\left(\frac{p_1}{p_0}\right)^i \left(\frac{p_2}{p_0}\right)^j \ldots \left(\frac{p_n}{p_0}\right)^l$ and therefore any polynomial in the n ratios $p_r : p_0$, is a symmetrical polynomial in the n roots $\alpha, \beta, \ldots, \lambda$, as direct substitution would show. It can be proved conversely that any symmetric polynomial in the roots is a polynomial in the n ratios $p_r : p_0$. The emphasis here is on the word polynomial, so that the functions considered are rational and integral, although if $n > 1$ each root $\alpha, \beta, \ldots$ separately is a very complicated irrational function of the coefficients p_r.

5. Gradient. Degree and Weight.

A polynomial function $\phi(\alpha)$ which is symmetric and homogeneous in the roots is a sum of such terms as

$$A\left(\frac{p_1}{p_0}\right)^i \left(\frac{p_2}{p_0}\right)^j \ldots \left(\frac{p_n}{p_0}\right)^l, \quad . \quad . \quad . \quad . \quad (21)$$

where A is a numerical coefficient. The factor p_r/p_0 is of degree r in the roots α. Hence for homogeneity in the roots, the expression

$$w = i + 2j + \ldots + nl \quad . \quad . \quad . \quad . \quad (22)$$

is constant for each such term. This number w is called the *weight* of the term. It is given by counting the total of suffixes of the p's in the term. Also it gives the degree of $\phi(\alpha)$ in the set $\alpha, \beta, \ldots, \lambda$, as relations (20) at once indicate.

After multiplying throughout by p_0^w, we can write of such a form,

$$p_0^w \phi(\alpha, \beta, \ldots, \lambda) = \psi(p_0, p_1, \ldots, p_n), \quad . \quad . \quad (23)$$

where both ϕ and ψ are homogeneous polynomials in their arguments. For p_0 no longer occurs in the denominator of any term.

Such a function ψ is isobaric, i.e. of the same weight w for each term. It is also homogeneous in the set p, since multiplying each of $p_0, p_1, \ldots, p_n$ by the same quantity t leaves $U_n = 0$, and therefore the roots $\alpha, \beta, \ldots$, unaltered. It is sometimes known as a *gradient*.

6. The Induced Linear Transformation of the Binary n-ic.

A binary form $U_n = (a_0, a_1, a_2, \ldots, a_n \between x, y)^n$ contains two *sets*, the set of coefficients

$$a = (a_0, a_1, \ldots, a_n) \quad . \quad . \quad . \quad . \quad . \quad . \quad (24)$$

and the set of variables x, y. It may seem a trivial remark, but it is one with far-reaching consequences, that a form is *linear* in its set of coefficients.

The transformation T (2) from x, y to x', y' is conveniently symbolized by an arrow. We write

$$T : x \rightarrow x', \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (25)$$

where x and x' do duty for x, y and x', y' respectively. For this reason it is preferable as a rule to use x_1, x_2 rather than x, y for two homogeneous variables.

By solving equations (2) we obtain the *inverse* transformation, written

$$T^{-1} : x' \rightarrow x. \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)$$

Provided the determinant $| M |$ of the transformation is not zero, this can always be done, even in the case of more than two variables.

The first important result of this theory is that *a form of order* n *remains a form of order* n *after linear transformation of its variables* x, y.

Thus, substituting for x, y in terms of x', y' we write

$$U_n = U(x, y) = U'(x', y'), \quad . \quad . \quad . \quad . \quad . \quad (27)$$

so that

$$a_0 x^n + \ldots + a_n y^n = a_0' x'^n + \ldots + a_n' y'^n. \quad . \quad . \quad (28)$$

This introduces a new coefficient set

$$a_0', a_1', \qquad , a_n', \tag{29}$$

such that each a_i' is a linear homogeneous function of the original set of coefficients a as in the particular case already worked out for the quadratic (6). The actual values of $a_0', a_1', \ldots, a_n'$ are best determined by Taylor's theorem. Thus

$$U(x, y) = U(l_1 x' + m_1 y', \quad l_2 x' + m_2 y'). \quad . \tag{30}$$

But U is homogeneous and of order n in x, y. So if $y' = tx'$, we have

$$U(x, y) = x'^n U(l_1 + m_1 t, \quad l_2 + m_2 t). \quad . \quad . \tag{31}$$

Expanding this as a function of two variables l_1, l_2, by Taylor's theorem, we get

$$x'^n \left\{ U(l_1, l_2) + t\left(m_1 \frac{\partial}{\partial l_1} + m_2 \frac{\partial}{\partial l_2}\right) U + \ldots + \frac{t^n}{n!}\left(m_1 \frac{\partial}{\partial l_1} + m_2 \frac{\partial}{\partial l_2}\right)^n U \right\},$$

in which m_1, m_2 are treated as constants, and U denotes $U(l_1, l_2)$. Again

$$\begin{aligned} U'(x', y') &= a_0' x'^n + \ldots + \binom{n}{r} a_r' x'^{n-r} y'^r + \ldots + a_n' y' \\ &= x'^n \left\{ a_0' + \ldots + a_n' t^n \right\}. \end{aligned}$$

Comparing coefficients of t^r in these two equivalent expansions we have the following result

$$\left.\begin{aligned} a_0' &= U(l_1, l_2) = (a_0, a_1, \ldots, a_n \between l_1, l_2)^n \\ a_1' &= \frac{1}{n}\left(m \frac{\partial}{\partial l}\right) U(l_1, l_2) \\ & \\ a_r' &= \frac{(n-r)!}{n!}\left(m \frac{\partial}{\partial l}\right)^r U(l_1, l_2) \\ & \\ a_n' &= \frac{1}{n!}\left(m \frac{\partial}{\partial l}\right)^n U(l_1, l_2) \end{aligned}\right\} \quad , \quad . \tag{32}$$

where $\left(m \frac{\partial}{\partial l}\right)$ denotes $m_1 \frac{\partial}{\partial l_1} + m_2 \frac{\partial}{\partial l_2}$.

As an illustration of this general set the reader is advised to write down the set of coefficients a_0', a_1', a_2', a_3' of a cubic after transformation T. It is important to appreciate that the relations giving $a_0', a_1' \ldots$ in terms of the old coefficients $a_0, a_1 \ldots$ are linear in these. We typify this by writing

$$T_{a'} : a' \rightarrow a,$$

and conversely

$$T_a : a \rightarrow a'$$

to denote the inverse transformation. This is a special case of the linear transformation of a set of $n + 1$ quantities. Since the coefficients of the a's are functions of l_1, m_1, l_2, m_2 and are thus completely determined by the coefficients of the original transformation

$$T : x \rightarrow x'$$

we call T_a or $T_{a'}$ an *induced linear transformation.*

Evidently there will be a close connexion between the determinant $\begin{vmatrix} l_1 m_1 \\ l_2 m_2 \end{vmatrix}$ of T and the determinant Δ of $T_{a'}$; in fact it is Δ raised to the power $\frac{1}{2}n(n + 1)$.

EXAMPLES

1. For the quadratic, prove

$$\Delta = \begin{vmatrix} l_1{}^2 & 2l_1 l_2 & l_2{}^2 \\ l_1 m_1 & l_1 m_2 + l_2 m_1 & l_2 m_2 \\ m_1{}^2 & 2m_1 m_2 & m_2{}^2 \end{vmatrix} = (l_1 m_2 - l_2 m_1)^3.$$

2. $$\begin{matrix} l_1{}^3, & 3l_1{}^2 l_2, & 3l_1 l_2{}^2, & l_2{}^3 \\ l_1{}^2 m_1, & 2l_1 l_2 m_1 + l_1{}^2 m_2, & m_1 l_2{}^2 + 2l_1 l_2 m_2, & l_2{}^2 m_2 \\ l_1 m_1{}^2, & l_2 m_1{}^2 + 2l_1 m_1 m_2, & 2m_1 m_2 l_2 + l_1 m_2{}^2, & l_2 m_2{}^2 \\ m_1{}^3, & 3m_1{}^2 m_2, & 3m_1 m_2{}^2, & \end{matrix} = (l_1 m_2 - l_2 m_1)^6.$$

Use $\text{col}_1 - \frac{l_1}{l_2} \text{col}_2 + \frac{l_1{}^2}{l_2{}^2} \text{col}_3 - \frac{l_1{}^3}{l_2{}^3} \text{col}_4$, $\text{col}_2 - 2\frac{l_1}{l_2} \text{col}_3 + 3\frac{l_1{}^2}{l_2{}^2} \text{col}_4$, &c.

3. Generalize the result.

7. Polar Forms.

In this last set (32) we have an example of the very important process known as *polarization.* It will be seen that the first coefficient a_0' is the binary n-ic

$$U(l_1, l_2) \equiv (a_0, a_1, \ldots, a_n \between l_1, l_2)^n,$$

with l_1, l_2 substituted for x, y. The other coefficients $a_1', a_2' \dots$ are the first, second, ..., nth polars respectively of a_0' with respect to m_1, m_2. These equations serve to define such polars.

In particular the last a_n' is the n-ic in m_1, m_2,

$$(a_0, a_1, \dots, a_n \between m_1, m_2)^n.$$

All intermediate coefficients a_r' $(0 < r < n)$ are examples of *double binary forms*: they possess double orders. Thus a_r' is of orders $(n - r, r)$ in the sets l_1, l_2 and m_1, m_2 respectively.

EXAMPLES

1. Write down the first and second polars of the quadratic $a_0x_1^2 + 2a_1x_1x_2 + a_2x_2^2$ with regard to the set y_1, y_2.

Ans. $a_0x_1y_1 + a_1(x_1y_2 + x_2y_1) + a_2x_2y_2$ and $a_0y_1^2 + 2a_1y_1y_2 + a_2y_2^2$.

2. Form the first and second polars of the cubic $a_0x_1^3 + 3a_1x_1^2x_2 + 3a_2x_1x_2^2 + a_3x_2^3$ with regard to the set y_1, y_2.

3. Find the rth polar of the binomial n-ic $ax_1^n + bx_2^n$ with regard to y_1, y_2.

Ans. $ax_1^{n-r}y_1^r + bx_2^{n-r}y_2^r$.

8. Formal Definition of Invariant.

If a binary form f *be changed by a linear transformation* T *into a new form* f′, *and a function* I *of the coefficients of* f′ *be equal to the same function of the coefficients of* f *multiplied by a factor depending solely on the transformation, then* I *is called an invariant of the binary form* f. *The form* f *is called the* **ground form.**

Let us write

$$I = I(a) = I(a_0, a_1, a_2, \dots, a_n)$$

to denote an invariant of the binary n-ic, whose coefficient set is

$$a = (a_0, a_1, a_2, \dots, a_n).$$

Further, let $T : x \rightarrow x'$ induce the transformation $T_a : a \rightarrow a'$, so that $I(a')$ would mean the same function of the $n + 1$ arguments

$$a_0', a_1', \dots, a_n'.$$

Then if

$$I(a') = \phi(l_1, l_2, m_1, m_2) I(a),$$

where ϕ depends solely on the four quantities l_1, l_2, m_1, m_2 and not on a or x, $I\ (a)$ is an invariant.

This definition follows from Boole's discovery.

Examples.—$\Delta = (a_0a_3 - a_1a_2)^2 - 4(a_0a_2 - a_1^2)(a_1a_3 - a_2^2)$ is an invariant of the binary cubic $(a_0, a_1, a_2, a_3 \between x, y)^3$.

For the quartic $(a_0, a_1, a_2, a_3, a_4 \between x, y)^4$, invariants are

$$I = a_0a_4 - 4a_1a_3 + 3a_2^2, \quad J = \begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix}$$

9. Simultaneous Invariants.

Let us reconsider our quadratic

$$U = ax^2 + 2hxy + by^2,$$

to which we adjoin a second,

$$V = Ax^2 + 2Hxy + By^2.$$

These in turn lead to a singly infinite system of quadratics typified by

$$U + \lambda V = (a + \lambda A)x^2 + 2(h + \lambda H)xy + (b + \lambda B)y^2.$$

Now consider the discriminant of $U + \lambda V$,

$$\begin{vmatrix} a + \lambda A & h + \lambda H \\ h + \lambda H & b + \lambda B \end{vmatrix}$$

It can be expanded in powers of λ and written

$$(ab - h^2) + \lambda(aB + bA - 2hH) + \lambda^2(AB - H^2).$$

But if a linear transformation T changes x to x', a to a', &c. we can write

$$U' + \lambda V' = U + \lambda V.$$

In particular

$$a' + \lambda A', \quad h' + \lambda H', \quad b' + \lambda B'$$

are the new coefficients of the quadratic. Hence, by (7), we have identically

$$\begin{matrix} a' + \lambda A' & h' + \lambda H' \\ h' + \lambda H' & b' + \lambda B' \end{matrix} = (l_1m_2 - l_2m_1)^2 \begin{matrix} a + \lambda A, & h + \lambda H \\ h + \lambda H, & b + \lambda B \end{matrix}$$

This is true for all values of a, h, b, A, H, B, λ and therefore we can equate coefficients of λ on each side. Thus

$$\begin{aligned} a'b' - h'^2 &= |M|^2 (ab - h^2), \\ a'B' + b'A' - 2h'H' &= |M|^2 (aB + bA - 2hH), \\ A'B' - H'^2 &= |M|^2 (AB - H^2). \end{aligned}$$

The first and third statements here tell us nothing new, but the second gives important information: it satisfies the characteristic invariant condition although it involves double as many coefficients a, h, b, A, H, B as the original quadratic. It introduces us to the new idea of a simultaneous invariant.

Definition of Simultaneous Invariant.—*If* (a), (b), . . . *denote the sets of coefficients* a_0, a_1, . . . , b_0, b_1, . . . *of different quantics, then* I (a, b, . . .) *is a simultaneous invariant of these quantics, provided*

$$I(a', b', \ldots) = \phi(l_1, l_2, m_1, m_2) I(a, b, \ldots)$$

identically for all values of the sets (a), (b), ...

10. The Aronhold Operator.

The above quadratic example leads to a general theorem due to Clebsch. Consider two p-ics

$$\begin{aligned} U &= (a_0, a_1, \ldots, a_p \between x, y)^p, \\ V &= (b_0, b_1, \ldots, b_p \between x, y)^p, \end{aligned} \qquad . \qquad (33)$$

and the *pencil* of p-ics given by

$$U + \lambda V = (a_0 + \lambda b_0)x^p + \ldots + \binom{p}{r}(a_r + \lambda b_r)x^{p-r}y^r + \ldots \\ + (a_p + \lambda b_p)y^p. \qquad . \quad . \quad . \quad . \quad (34)$$

Here is an example of the addition theorem of linear sets. In the matrix notation we could write the coefficient set of $U + \lambda V$

$$[a + \lambda b] = [a] + \lambda[b]. \qquad . \quad . \quad . \quad . \quad (35)$$

Let $T : x \rightarrow x'$ be a linear transformation changing U to U', V to V', and therefore giving a_i' linearly in terms of set $[a]$, b_i' linearly in terms of set $[b]$. Hence

$$[a' + \lambda b'] = [a'] + \lambda[b']. \qquad (36)$$

Now suppose $I(a_0, a_1, \ldots, a_n)$, written $I(a)$, to be an invariant of U, so that

$$I(a') = \phi \,.\, I(a):$$

then, from the definition,

$$I(a' + \lambda b') = \phi \,.\, I(a + \lambda b)$$

identically for all values of λ. Expand both sides by Taylor's theorem and equate powers of λ. The coefficient of λ on each side gives

$$\left(b_0' \frac{\partial}{\partial a_0'} + \ldots + b_p' \frac{\partial}{\partial a_p'}\right) I(a_0', a_1', \ldots, a_p')$$
$$= \phi \,.\, \left(b_0 \frac{\partial}{\partial a_0} + \ldots + b_p \frac{\partial}{\partial a_p}\right) I(a_0, a_1, \ldots, a_p). \quad (37)$$

This is usually written in the contracted notation

$$\left(b' \frac{\partial}{\partial a'}\right) I(a') = \phi \,.\, \left(b \frac{\partial}{\partial a}\right) I(a). \quad . \quad . \quad . \quad (38)$$

Likewise the coefficient of λ^r gives

$$\left(b' \frac{\partial}{\partial a'}\right)^r I(a') = \phi \,.\, \left(b \frac{\partial}{\partial a}\right)^r I(a) \quad . \quad . \quad . \quad (39)$$

where the arguments $b_0, b_1, \ldots, b_p$ are independent of $a_0, a_1, \ldots, a_p$, and so must not be differentiated in the course of the work. But these last results yield functions of both sets $[a]$, $[b]$ which satisfy the invariant condition.

We conclude that the operator

$$\left(b \frac{\partial}{\partial a}\right) = \sum_{i=0}^{p} b_i \frac{\partial}{\partial a_i} \quad . \quad . \quad . \quad . \quad . \quad . \quad (40)$$

applied to an invariant involving $a_0, a_1, \ldots, a_p$ produces an invariant. For this reason it is called an invariant process. In particular from a rational integral homogeneous invariant of degree q in the set $a_0, a_1, \ldots, a_p$, it produces $q - 1$ simultaneous invariants involving both sets $[a]$ and $[b]$.

The name Aronhold operator is sometimes given to $\left(b \frac{\partial}{\partial a}\right)$, after one of the founders of the theory.

Definition of Invariant Process.—*If the effect of a process* R *applied to a function of the original coefficients* a *is the same as* *that of the process applied to the corresponding function of the new coefficients, save for the factor* ϕ, *then* R *is called an invariant process.*

Formula (38) is an example of this; and it will appear that polarization in general is an invariant process.

11. Multilinear Invariants.

The invariant

$$aB + bA - 2hH,$$

which is a bilinear form in the sets a, h, b and A, H, B, could be written

$$\left(A\frac{\partial}{\partial a} + H\frac{\partial}{\partial h} + B\frac{\partial}{\partial b}\right)(ab - h^2),$$

as an illustration of the Aronhold process. Suppose, however, we had a homogeneous invariant of degree q in a set $a_0, a_1, \ldots, a_p$. We write it

$$I = I(a),$$

and proceed to operate with $\left(b\frac{\partial}{\partial a}\right)$. It produces an invariant homogeneous in both $[a]$ and $[b]$ of degrees $q - 1$ and 1 respectively. We could write it I_1, so that

$$I_1(a', b') = \phi \, . \, I_1(a, b)$$

identically for all values of $[a]$, $[b]$. Now we choose a third quantic

$$(c_0, c_1, \ldots, c_p \between x, y)^p$$

and operate with $\left(c\frac{\partial}{\partial a}\right)$ on I, treating both $[c]$ and $[b]$ as independent of the set $a_0, a_1, \ldots, a_p$. The result as before is an invariant, this time linear in $[b]$, linear in $[c]$ and of degree $q - 2$ in $[a]$. Thus

$$I_2(a, b, c) = \left(c\frac{\partial}{\partial a}\right) I_1(a, b) = \left(c\frac{\partial}{\partial a}\right)\left(b\frac{\partial}{\partial a}\right) I(a).$$

Proceeding in this way to q operations involving q different quantics all of order p, we finally deduce a multilinear invariant

$$I_q(b, c, \ldots, k)$$

involving q sets of coefficients $[b], \ldots, [k]$ of q different p-ics.

Mathematics is a beautiful subject.

Again, since the coefficients $b_0, \ldots, k_p$ have been taken quite generally as independent of $a_0, \ldots, a_p$, *the Aronhold operators are commutative.* In fact

$$\left(b\frac{\partial}{\partial a}\right)\left(c\frac{\partial}{\partial a}\right)f(a) = \sum_{i,j=0}^{n}\sum b_i c_j \frac{\partial^2}{\partial a_i \partial a_j} f(a) = \left(c\frac{\partial}{\partial a}\right)\left(b\frac{\partial}{\partial a}\right)f(a).$$

It follows that I_q is symmetrical in $[b]$, $[c]$ and therefore in all of $b, c, \ldots, k$. This means that the $q!$ arrangements of $b, c, \ldots, k$ are equivalent, so that

$$I_q(b, c, \ldots, k) = I_q(c, b, \ldots, k) = \ldots.$$

Further, by Euler's theorem for homogeneous forms

$$\left(a\frac{\partial}{\partial a}\right)I(a) = \sum_{i=0}^{p} a_i \frac{\partial I}{\partial a_i} = qI(a).$$

Hence we infer that the result of putting $[b] = [a]$ in $\left(b\frac{\partial}{\partial a}\right)I(a)$ is merely to multiply $I(a)$ by q. Thus we have a useful theorem:

Every invariant of a binary p-*ic* f, *homogeneous and of degree* q *in the coefficients of* f, *may be regarded as a special case of an invariant at once linear and symmetrical in the* q *sets of coefficients of* q *binary* p-*ics.*

EXAMPLES

1. Form an invariant of two quartics $(a_0, a_1, a_2, a_3, a_4 \between x, y)^4$ and $(b_0, b_1, b_2, b_3, b_4 \between x, y)^4$ linear in each.

Ans. $a_0b_4 - 4a_1b_3 + 6a_2b_2 - 4a_3b_1 + a_4b_0$.

2. Form an invariant linear in a quartic and of degree two in a quadratic.

Hint.—Consider the square of the quadratic as a quartic.

12. Covariants.

Let us once more return to the binary quadratics

$$\begin{aligned} U &= ax^2 + 2hxy + by^2 \\ V &= Ax^2 + 2Hxy + By^2 \end{aligned} \quad . \quad . \quad . \quad . \quad (41)$$

and form their Jacobian (§6, p. 124) or functional determinant,

$$J = \begin{vmatrix} \frac{\partial U}{\partial x} & \frac{\partial V}{\partial x} \\ \frac{\partial U}{\partial y} & \frac{\partial V}{\partial y} \end{vmatrix} = \frac{\partial(U, V)}{\partial(x, y)} = 4(U, V) \quad . \quad . \quad . \quad (42)$$

introducing various useful notations. This is

$$\begin{matrix} ax+hy & Ax+Hy \\ hx+by & Hx+By \end{matrix} = 4(\alpha x^2+2\beta xy+\gamma y^2), \quad (43)$$

say, which is another quadratic. Similarly, with the accented notation for the effect of the transformation $T: x \longrightarrow x'$, we have

$$4(\alpha' x'^2+2\beta' x' y'+\gamma' y'^2) = 4\begin{vmatrix} a'x'+h'y', & A'x'+H'y' \\ h'x'+b'y', & H'x'+B'y' \end{vmatrix}$$

$$= J' = \begin{vmatrix} \dfrac{\partial U'}{\partial x'}, & \dfrac{\partial V'}{\partial x'} \\ \dfrac{\partial U'}{\partial y'}, & \dfrac{\partial V'}{\partial y'} \end{vmatrix} \quad (44)$$

But $U(x, y) = U'(x', y')$, $V(x, y) = V'(x', y')$. Accordingly

$$\frac{\partial U'}{\partial x'} = \frac{\partial U}{\partial x}\,\frac{\partial x}{\partial x'} + \frac{\partial U}{\partial y}\,\frac{\partial y}{\partial x'} = l_1\frac{\partial U}{\partial x} + l_2\frac{\partial U}{\partial y},$$

and
$$\frac{\partial U'}{\partial y'} = m_1\frac{\partial U}{\partial x} + m_2\frac{\partial U}{\partial y}.$$

Similarly for V. Substituting in (44) we find

$$J' = \begin{matrix} l_1\dfrac{\partial U}{\partial x} + l_2\dfrac{\partial U}{\partial y}, & l_1\dfrac{\partial V}{\partial x} + l_2\dfrac{\partial V}{\partial y} \\ m_1\dfrac{\partial U}{\partial x} + m_2\dfrac{\partial U}{\partial y}, & m_1\dfrac{\partial V}{\partial x} + m_2\dfrac{\partial V}{\partial y} \end{matrix}$$

$$\begin{matrix} l_1 & l_2 \\ m_1 & m_2 \end{matrix}\Bigg| \; \begin{vmatrix} \dfrac{\partial U}{\partial x} & \dfrac{\partial V}{\partial x} \\ \dfrac{\partial U}{\partial y} & \dfrac{\partial V}{\partial y} \end{vmatrix} = |M|\,J.$$

This introduces us to a covariant of the given forms U, V: namely a function of their coefficients, *and their variables* x, y, which maintains itself after linear transformation, but for a factor depending solely on the linear transformation.

Definition of Covariant:—*In the notation already adopted, a* *function* C *of sets* [a], [b], . . . *of coefficients of different quantics whose variables are* x_1, x_2 *is a co-variant, if*

$$C(a', b', \ldots, x') = \phi(l_1, l_2, m_1, m_2)C(a, b, \ldots, x)$$

identically for all values of $a_0, a_1, \ldots, b_0, b_1, \ldots, x_1, x_2$.

In the above example C is the Jacobian, which is a function of eight arguments a, h, b, A, H, B, x, y. But in detail it is bilinear in the sets a, h, b and A, H, B, while being quadratic in the set of variables x, y.

In what follows our chief concern is with rational integral homogeneous invariants and covariants.

13. Relation between Linear Forms and Covariants.

The simplest quantic to deal with is the linear form

$$E = e_1x_1 + e_2x_2. \quad . \quad . \quad . \quad . \quad . \quad (45)$$

Presumably invariants exist involving this and other forms. For example, it is easy to verify that

$$a_0e_2^2 - 2a_1e_1e_2 + a_2e_1^2 \quad . \quad . \quad . \quad . \quad . \quad (46)$$

is an invariant of E and the quadratic $a_0x_1^2 + 2a_1x_1x_2 + a_2x_2^2$. In fact if $a \rightarrow a'$, $e \rightarrow e'$ denote the linear transformations, we have, by (32),

$$\begin{aligned} e_1' &= l_1e_1 + l_2e_2 \\ e_2' &= m_1e_1 + m_2e_2 \end{aligned} \quad \text{or} \quad \begin{aligned} e_1\,|M| &= m_2e_1' - l_2e_2' \\ -e_2\,|M| &= m_1e_1' - l_1e_2'. \end{aligned} \quad (47)$$

But

$$a_0x_1^2 + 2a_1x_1x_2 + a_2x_2^2 = a_0'x_1'^2 + 2a_1'x_1'x_2' + a_2'x_2'^2 \quad (48)$$

identically for all values of x_1, x_2. So, in particular, let

$$x_1' = -e_2', \quad x_2' = +e_1',$$

then

$$\begin{aligned} x_1 &= l_1x_1' + m_1x_2' = -l_1e_2' + m_1e_1' = -e_2\,|M| \\ x_2 &= l_2x_1' + m_2x_2' = -l_2e_2' + m_2e_1' = e_1\,|M| \end{aligned} \quad (49)$$

Substitute in (48): then

$$a_0'e_2'^2 - 2a_1'e_1'e_2' + a_2'e_1'^2 = |M|^2(a_0e_2^2 - 2a_1e_1e_2 + a_2e_1^2), \quad (50)$$

which exhibits the invariant property.

This feature is true in general; indeed any polynomial covariant can be looked upon as an invariant of the linear form $e_1x_1+e_2x_2$. For if $C(a_i, x_1, x_2)$ denotes the covariant in question, a_i standing for the coefficients of the ground form (or forms), then by the characteristic property

$$C(a_i', x_1', x_2')=\phi\, C(a_i, x_1, x_2). \quad . \quad . \quad . \quad (51)$$

Since $x_1=ty_1$, $x_2=ty_2$ is a particular case of the linear transformation, then

$$C(a_i', y_1, y_2)=\phi\, C(a_i, ty_1, ty_2), \quad . \quad . \quad . \quad (52)$$

where ϕ depends solely on t. This implies that C is homogeneous in the variables y_1, y_2 as the contrary assumption at once is seen to be impossible when applied to (52).

Hence (51) is homogeneous in x_1, x_2 and, let us say, of order ϖ. Using (49) this at once yields

$$\begin{aligned}C(a_i', -e_2', e_1')&=\phi C(a_i, -e_2\,|\,M\,|, e_1\,|\,M\,|)\\&=\phi\,|\,M\,|^{\tilde{\omega}}\cdot C(a_i, -e_2, e_1).\end{aligned}$$

This shows that e_1, e_2 enter the function C precisely as e_1', e_2' do, so that the function $C(a_i, -e_2, e_1)$ is an invariant of the original ground forms together with the linear form $e_1x_1+e_2x_2$.

A good physicist is a better mathematician.

CHAPTER IX

THE GENERAL LINEAR TRANSFORMATION

1. Cogredience and Contragredience.

The binary forms have served to introduce certain ideas which can easily be generalized. We shall now be concerned with forms in n variables

$$x = \{x_1, x_2, x_3, \ldots, x_n\},$$

which undergo a linear transformation

$$T_x: \quad \begin{array}{l} x_1 = \xi_1 x_1' + \eta_1 x_2' + \ldots + \omega_1 x_n', \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_i = \xi_i x_1' + \eta_i x_2' + \ldots + \omega_i x_n', \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_n = \xi_n x_1' + \eta_n x_2' + \ldots + \omega_n x_n', \end{array} \tag{1}$$

or $T_x : x \longrightarrow x'$. Let M denote the square matrix of coefficients $\xi_1, \ldots, \omega_n$, and $|M|$ its determinant, so that

$$M = \begin{bmatrix} \xi_1 & \eta_1 & \ldots & \omega_1 \\ & & & \\ \xi_n & \eta_n & \ldots & \omega_n \end{bmatrix}, \quad |M| = (\xi\eta \ldots \omega), \tag{2}$$

which must not vanish. The variables and coefficients may be real or complex numbers.

Let the co-factor of ξ_i in $|M|$ be X_i, and the reciprocal of $|M|$ be

$$\frac{1}{|M|} : \quad \begin{vmatrix} \xi^1 & \xi^2 & \ldots & \xi^n \\ & & & \\ & & & \omega^n \end{vmatrix} \tag{3}$$

so that $\xi^i = X_i / |M|$, &c.

If we multiply rows of T_x by $\xi^1, \xi^2, \dots, \xi^n$ respectively and add, we have

$$x_1' = \xi^1 x_1 + \xi^2 x_2 + \dots + \xi^n x_n,$$

and similarly

$$\begin{aligned} x_2' &= \eta^1 x_1 + \eta^2 x_2 + \dots + \eta^n x_n \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ x_n' &= \omega^1 x_1 + \omega^2 x_2 + \dots + \omega^n x_n. \end{aligned} \qquad (4)$$

This set of equations, which forms the inverse transformation of T_x can be written

$$T_x^{-1} : x' \to x. \qquad (5)$$

It exists provided $| M |$ is non-zero.

Now suppose we have a linear form

$$(u \mid x) = u_1 x_1 + u_2 x_2 + \dots + u_n x_n; \qquad (6)$$

let us consider the effect of the linear transformation (1) upon $(u \mid x)$. Manifestly it gives a form linear in the new set $[x']$; and if we denote the new coefficients by $u_1', u_2', \dots, u_n'$, we have

$$\begin{aligned} (u \mid x) = (u_1 \xi_1 + u_2 \xi_2 + \dots + u_n \xi_n) x_1' + \dots \\ + (u_1 \omega_1 + u_2 \omega_2 + \dots + u_n \omega_n) x_n' = (u' \mid x'), \end{aligned} \qquad (7)$$

where the new coefficients u' are linear functions of the old. Accordingly we write

$$T_u^{-1} : \quad \begin{aligned} u_1' &= \xi_1 u_1 + \xi_2 u_2 + \dots + \xi_n u_n, \\ u_2' &= \eta_1 u_1 + \eta_2 u_2 + \dots + \eta_n u_n, \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ u_n' &= \omega_1 u_1 + \omega_2 u_2 + \dots + \omega_n u_n. \end{aligned} \qquad (8)$$

Here we have another instance of an induced linear transformation $u' \to u$, where, it will be noticed, the coefficient matrix is the transposed of matrix M in (1). We therefore write it M'. By solving this set (8) we obtain the direct transformation $u \to u'$ which shall be denoted by T_u, so that its inverse T_u^{-1} denotes (8). Thus

$$T_u : \quad \begin{aligned} u_1 &= \xi^1 u_1' + \eta^1 u_2' + \dots + \omega^1 u_n', \\ &\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ u_n &= \xi^n u_1' + \eta^n u_2' + \dots + \omega^n u_n'. \end{aligned} \qquad (9)$$

In this way we arrive at four transformations T_x, T_x^{-1}, T_u, T_u^{-1} as stated in (1), (4), (9), (8), whose matrices are M, M^{-1}, $(M')^{-1}$, M' respectively. When two such sets $[x]$ and $[u]$ undergo such transformations, (1) and (9), they are called[1] *contragredient* sets, and the same name is given to the corresponding transformations T_x, T_u. Further if $y_1, y_2, \ldots, y_n$ is another set of variables which undergoes the same transformation as $x_1, x_2, \ldots, x_n$, namely

$$T_x : y \rightarrow y', \quad \left.\begin{array}{c} y = \xi_i y_1' + \eta_i y_2' + \ldots + \omega_i y_n' \\ i = 1, 2, \ldots, n, \end{array}\right\} \quad (10)$$

then the sets $[x]$ and $[y]$ are called *cogredient.*

The simplest formal definition of cogrediency and contragrediency is to take them as follows:

Two sets of n *variables* [x] *and* [y] *are cogredient if a linear transformation of matrix* M *for* $x \rightarrow x'$ *induces the transformation* $y \rightarrow y'$ *with the same matrix. Two sets* [x] *and* [u] *are contragredient if, when* $x \rightarrow x'$ *and* $u \rightarrow u'$, *the inner product* u_x *remains an absolute invariant, namely*

$$u_1 x_1 + \ldots + u_n x_n = u_1' x_1' + \ldots + u_n' x_n', \quad . \quad (11)$$

or simply $$(u \mid x) = (u' \mid x').$$

Starting from this fundamental condition, which must hold identically, we can at once deduce equations (8) from equations (1) by substituting in (11); or conversely.

2. Linear Transformations in Matrix Notation.

Let U denote the single-row matrix

$$[u_1, u_2, \ldots, u_n] \quad . \quad . \quad . \quad . \quad . \quad (12)$$

and X the single *column* matrix whose transposed is

$$X' = [x_1, x_2, \ldots, x_n]. \quad . \quad . \quad . \quad . \quad (13)$$

Let $$\overline{X}' = [x_1', x_2', \ldots, x_n']. \quad . \quad . \quad . \quad (14)$$

Then *the general linear transformation* (1) *is a direct example of the product of matrices*, and can be written

$$T_x: \qquad X = M\overline{X}. \quad . \quad . \quad . \quad . \quad . \quad (15)$$

---

[1] Sylvester first developed this theory, and gave these names to the sets $[u]$, $[x]$. Cf. *Cambridge and Dublin Mathematical Journal,* VI, VII, VIII, IX (1851–4).

as is immediately apparent when it is written in full. Next let cogredient sets be denoted by single-column matrices X, Y, $Z \dots$. If they transform to $\bar{X}$, $\bar{Y}$, $\bar{Z}, \dots$, then by definition of cogredient sets

$$X = M\bar{X}, \quad Y = M\bar{Y}, \quad Z = M\bar{Z}. \quad . \quad . \quad (16)$$

We deduce by fore multiplication with M^{-1} that

$$M^{-1}X = \bar{X}, \quad M^{-1}Y = \bar{Y}, \quad M^{-1}Z = \bar{Z}. \quad . \quad (17)$$

Again, by (11), the contragredient sets U, X satisfy the identical condition between two inner products, which in matrix notation is

$$UX = \bar{U}\bar{X}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

Hence by (15), $$UM\bar{X} = \bar{U}\bar{X},$$

identically, so that

$$UM = \bar{U}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

which is the matrix equation for (8). Solving this we have

$$U = \bar{U}M^{-1}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (20)$$

which is the set of equations (9). By (9), p. 70, we may transpose these last results to

$$\bar{U}' = M'U' \quad U' = (M^{-1})'\bar{U}'.$$

giving the same actual equations when written in full.

If V is a set cogredient with U, then by (19)

$$\bar{V} = VM,$$

whence $$\bar{V}\bar{X} = VM\bar{X} = VX$$

so that V is contragredient with X.

Thus we arrive at the conception of a number of matrices or vectors of the first kind X, Y, $Z, \dots$, and a number of matrices or vectors of the second kind U, V, $W, \dots$, such that vectors of the same kind are cogredient with each other and vectors of different kinds are contragredient. All such matrices have rank unity (or else zero), for they each consist of a single row or column. Sometimes they are called tensors of the first rank, or order (cf. p. 91).

The two chief applications of this co- and contra-gredience are first in geometry and secondly in analysis, as the following merely preliminary examples are designed to show.

EXAMPLES

1. If $n = 3$, X may represent a point whose homogeneous co-ordinates referred to a triangle in a plane are x_1, x_2, x_3. If u_1, u_2, u_3 are homogeneous line co-ordinates the equation of a straight line is

$$u_1x_1 + u_2x_2 + u_3x_3 = 0.$$

If a new triangle of reference is chosen such that x_1', x_2', x_3' are the co-ordinates of the same point as before, and u_1', u_2', u_3' those of the same line as before, then the characteristic contragredient condition $u_x = u'_{x'}$ is satisfied. Hence in a plane, *homogeneous line and point co-ordinates are contragredient sets.*

2. Sets of co-ordinates of coplanar points $X, Y, Z \ldots$ are cogredient.

3. Sets of co-ordinates of coplanar lines $U, V, W \ldots$ are cogredient.

4. Points and planes in threefold space are contragredient. [Here $n = 4$.]

5. If φ is a function of 2 variables x, y and $x = r\cos\theta$, $y = r\sin\theta$, prove that the set $[dx, dy]$ is contragredient to $\left[\frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial y}\right]$ relative to the *linear* transformation of differentials from $[dx, dy]$ to $[dr, d\theta]$.

For let $$\varphi(x, y) = \varphi'(r, \theta) = \varphi.$$

Then $$d\varphi = \frac{\partial\varphi}{\partial x}\,dx + \frac{\partial\varphi}{\partial y}\,dy = \frac{\partial\varphi'}{\partial r}\,dr + \frac{\partial\varphi}{\partial\theta}\,d\theta.$$

Also $$dx = \cos\theta\ dr - r\sin\theta\ d$$
$$dy = \sin\theta\ dr + r\cos\theta\ d$$

These last give a linear transformation, of modulus r, for the differentials dx, dy in terms of $dr, d\theta$. The proof is now immediate.

6. Write down the induced linear transformation of $\frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial y}$.

$$\left[\frac{\partial\varphi'}{\partial r} = \cos\theta\,\frac{\partial\varphi}{\partial x} + \sin\theta\,\frac{\partial\varphi}{\partial y},\quad \frac{\partial\varphi'}{\partial\theta} = -r\sin\theta\,\frac{\partial\varphi}{\partial x} + r\cos\theta\,\frac{\partial\varphi}{\partial y}.\right]$$

7. If $x = x(p, q)$, $y = y(p, q)$ prove that $[dx, dy]$ and $\left[\frac{\partial\varphi}{\partial x}, \frac{\partial\varphi}{\partial y}\right]$ are still contragredient, relative to the linear transformation $[dx, dy] \longrightarrow [dp, dq]$, provided the Jacobian $\frac{\partial(x, y)}{\partial(p, q)} \equiv \frac{\partial x}{\partial p}\frac{\partial y}{\partial q} - \frac{\partial x}{\partial q}\frac{\partial y}{\partial p}$ neither vanishes nor is infinite.

[This Jacobian is the determinant of the linear transformation.]

8. Generalize this for n variables, $x_1, x_2, \ldots, x_n$, proving that $[dx_1, dx_2, \ldots, dx_n]$ and $\left[\frac{\partial\varphi}{\partial x_1}, \frac{\partial\varphi}{\partial x_2}, \ldots, \frac{\partial\varphi}{\partial x_n}\right]$ are contragredient sets.

3. Orthogonal Transformations and Matrices.

We gain a clearer idea of cogredience and contragredience of sets of variables by considering a particular case in which the distinction breaks down. It is called the orthogonal transformation, an example of which has already been considered (§2, p. 130). But the general orthogonal case is most fruitfully developed by starting with the characteristic property of contragredience of two sets (u) and (x) and seeking to make it hold of a single set (x) with itself.

Let $X = AY$ be the linear transformation with non-singular matrix A for a column set $X = \{x_1, x_2, \ldots, x_n\}$ in terms of another such set $Y = \{y_1, y_2, \ldots, y_n\}$. Then by transposition

$$X' = [x_1, x_2, \ldots, x_n], \qquad Y' = [y_1, y_2, \ldots, y_n] \quad (21)$$

and, for the inner products,

$$\begin{aligned} X'X &= x_1^2 + x_2^2 + \ldots + x_n^2 = (x \mid x) \\ Y'Y &= y_1^2 + y_2^2 + \ldots + y_n^2 = (y \mid y). \end{aligned} \quad (22)$$

Definition of Orthogonal Transformation.—*The homogeneous linear transformation from* $x_1, x_2, \ldots, x_n$ *to* $y_1, y_2, \ldots, y_n$ *is orthogonal if the condition*

$$x_1^2 + x_2^2 + \ldots + x_n^2 = y_1^2 + y_2^2 + \ldots + y_n^2 \quad (23)$$

is identically satisfied by performing this transformation.

This condition can be written in either of the equivalent forms

$$X'X = Y'Y, \qquad (x \mid x) = (y \mid y). \quad . \quad . \quad (24)$$

To fix our ideas, let the typical case when $n = 3$ be taken. Then we suppose that the following transformation is orthogonal:

$$\begin{aligned} x_1 &= a_1y_1 + b_1y_2 + c_1y_3, \\ x_2 &= a_2y_1 + b_2y_2 + c_2y_3, \\ x_3 &= a_3y_1 + b_3y_2 + c_3y_3, \end{aligned} \quad (25)$$

also written

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} \quad (26)$$

or simply

$$X = AY.$$

Thus, if when $n = 3$ the substitutions (25) are made in (23), the result is a quadratic condition involving terms in y_1^2, y_2^2, y_3^2, y_2y_3, y_3y_1, y_1y_2. Since this is true for all values of y_1, y_2, y_3 the coefficients of these six quadratic terms must vanish identically. This gives

$$\begin{aligned} a_1^2 + a_2^2 + a_3^2 = 1, \quad & b_1c_1 + b_2c_2 + b_3c_3 = 0, \\ b_1^2 + b_2^2 + b_3^2 = 1, \quad & c_1a_1 + c_2a_2 + c_3a_3 = 0, \\ c_1^2 + c_2^2 + c_3^2 = 1, \quad & a_1b_1 + a_2b_2 + a_3b_3 = 0, \end{aligned} \qquad (27)$$

which is completely specified by the matrix equation

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} = \begin{bmatrix} 1 & . & . \\ . & 1 & . \\ . & . & 1 \end{bmatrix} \qquad (28)$$

or simply

$$A'A = I, \qquad (29)$$

which is also true for all values of n.

This last result characterizes the *orthogonal matrix*, namely *the product of an orthogonal matrix* A *and its transposed* A′ *is the unit matrix.*

Further if $A'A = I$, then the product of the corresponding determinants gives

$$|A'|\;|A| \qquad |A|^2 = 1 \qquad (30)$$

so that the determinant $|A|$ is ± 1. In either case the inverse A^{-1} exists, for the matrix is non-singular.

Now

$$AA'A = A(A'A) = AI = A.$$

Hence by after-multiplication

$$AA'AA^{-1} = AA^{-1} = I,$$

so that $AA' = I$: hence *an orthogonal matrix commutes with its transposed*, and

$$A'A = AA' = I. \quad . \quad . \quad . \quad . \quad . \quad . \quad (31)$$

Conversely, if $A'A = I$, we deduce the original property (24) of the sets X, Y. For if

$$X = AY, \qquad X' = Y'A',$$

then $$X'X = (Y'A')(AY) = Y'A'AY = Y'IY = Y'Y,$$

which exhibits the required property (24).

If we expand the result (31) to its full implication, when $n = 3$, we obtain the six equations (27), together with a further six due to transposition. Thus we interchange the sets of suffixes and letters in (27) and obtain

$$\begin{array}{ll} a_1^2 + b_1^2 + c_1^2 = 1 & a_2a_3 + b_2b_3 + c_2c_3 = 0, \\ a_2^2 + b_2^2 + c_2^2 = 1 & a_3a_1 + b_3b_1 + c_3c_1 = 0, \\ a_3^2 + b_3^2 + c_3^2 = 1 & a_1a_2 + b_1b_2 + c_1c_2 = 0. \end{array} \qquad (32)$$

Similarly, for n rows and columns the conditions (32) imply conditions (27), and conversely. Counting the number of such equations (27) in general, the number of necessary and sufficient conditions for A to be orthogonal is

$$n + \binom{n}{2} = \frac{n(n+1)}{2}.$$

Since A has n^2 elements, there are therefore $\dfrac{n(n-1)}{2}$ arbitrary constants involved in an orthogonal matrix.

A simple way to remember the conditions (27) or (32) is this: *the inner product of two different rows* (*or two different columns*) *of the orthogonal matrix is zero*; *that of a row or column with itself is unity.*

It will now be seen that the binary illustration of §2, p. 130, fits in with this general treatment of the orthogonal matrix. Further, since $|A| = \pm 1$ the notion of right- and left-handedness can also be attached to a general orthogonal transformation.

EXAMPLES

1. If A is an orthogonal matrix, so is A': and so also are A^{-1}, A'^{-1}.

2. $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ is orthogonal.

3. If $x = x' \cos\theta - y' \sin\theta$, $y = x' \sin\theta + y' \cos\theta$, the consequent orthogonal matrix characterizes the rotation of rectangular Cartesian axes through an angle θ. Its determinant is unity.

4. Show that $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$ characterizes a change of axes obtained by rotation through an angle θ followed by reversal of the axis of y'. Its determinant is -1.

5. The ternary ($n = 3$) orthogonal transformation, when $|A| = 1$, characterizes a change of rectangular Cartesian axes with fixed origin, obtained by suitable rotation.

{The matrix A gives direction cosines of old axes referred to new, or vice versa.}

6. If $|A| = -1$ the change of axes involves reflexion together with rotation.

7. *An orthogonal transformation, when* $|A| = 1$, *also characterizes a movement of a rigid body about a fixed pivot.*

For if P, Q, X, Y are column matrices, such that $P = AX$, $Q = AY$, then $P'P = X'X$, $Q'Q = Y'Y$, $P'Q = X'Y$. And if these are interpreted geometrically for rectangular Cartesian axes, $P'P$ means the square of the distance of a point P from the origin, while $P'Q$ gives $OP . OQ \cos POQ$. Hence the matrix conditions show that triangles POQ, XOY are congruent.

4. Cayley's Determination of the Orthogonal Matrix all whose Latent Roots differ from Negative Unity.

Let S be a general skew symmetric matrix and L be the sum of S and the unit matrix I, so that if $n = 3$, we write

$$S = \begin{matrix} & & -b \\ -c & & a \\ & -a & \end{matrix} \qquad L = \begin{bmatrix} 1 & c & -b \\ -c & 1 & a \\ & & 1 \end{bmatrix} \qquad (33)$$

and in general

$$S = -S', \quad L = I + S, \quad L' = I + S' = I - S. \quad (34)$$

Also let X, Y, Z be the column matrices of three sets of variables

$$\{x_1, x_2, \dots, x_n\}, \quad \{y_1, y_2, \dots, y_n\}, \quad \{z_1, z_2, \dots, z_n\}.$$

Then *if* $X = LZ$ *and* $Y = L'Z$, *the direct transformation from* X *to* Y *is orthogonal, and yields an orthogonal matrix, with no latent root equal to* -1 *and with the* $\frac{1}{2}n(n-1)$ *elements of a general skew symmetric matrix* S *for its arbitrary constants.*

In effect this is the theorem of Cayley[1]; nor is it difficult to prove. To fix our ideas let the conditions be written in full, when $n = 3$; namely $X = LZ$, $Y = L'Z$ become

$$\begin{array}{ll} x_1 = z_1 + cz_2 - bz_3 & y_1 = z_1 - cz_2 + bz_3 \\ x_2 = -cz_1 + z_2 + az_3 & y_2 = cz_1 + z_2 - az_3 . \\ x_3 = bz_1 - az_2 + z_3 & y_3 = -bz_1 + az_2 + z_3 \end{array} \qquad (35)$$

According to this theorem the effect of solving for the set z in terms of x and substituting in the set of equations for y, will give us an orthogonal transformation from y to x. In fact since $X = LZ$, therefore $X' = Z'L'$, so that $X'X = Z'L'LZ$.

Similarly $$Y'Y = Z'LL'Z.$$

But $$LL' = (I + S)(I + S') = (I + S)(I - S)$$

and $$L'L = (I + S')(I + S) = (I - S)(I + S).$$

Hence $X'X = Y'Y$, which proves the orthogonal property.

Further, we have $Y = L'Z$, so $Z = L'^{-1}Y$, provided L' is non-singular. Hence

$$X = LZ = LL'^{-1}Y. \quad . \quad . \quad . \quad . \quad (36)$$

Since L commutes with L', it commutes with L'^{-1}; hence LL'^{-1} can be written without ambiguity as

$$\frac{L}{L'} = \frac{I + S}{I - S}. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (37)$$

Thus the matrix A, which can be put into the form

$$\frac{I + S}{I - S},$$

where S is an arbitrary skew symmetric matrix, is orthogonal.

There is no difficulty in calculating A, since $(I + S)(I - S)^{-1}$ is given by L and the inverse of L'. In the case when

$$S = \begin{array}{ccc} & & -b \\ -c & . & a \\ b & -a & \end{array}$$

---

[1] *Crelle*, **32** (1846), 119–123; *Collected Works*, **1**, 332–336.

we find

$$A = \begin{bmatrix} \dfrac{1+a^2-b^2-c^2}{1+a^2+b^2+c^2}, & \dfrac{2(ab+c)}{1+a^2+b^2+c^2}, & \dfrac{2(ac-b)}{1+a^2+b^2+c^2} \\[2ex] \dfrac{2(ab-c)}{1+a^2+b^2+c^2}, & \dfrac{1-a^2+b^2-c^2}{1+a^2+b^2+c^2}, & \dfrac{2(bc+a)}{1+a^2+b^2+c^2} \\[2ex] \dfrac{2(ac+b)}{1+a^2+b^2+c^2}, & \dfrac{2(bc-a)}{1+a^2+b^2+c^2}, & \dfrac{1-a^2-b^2+c^2}{1+a^2+b^2+c^2} \end{bmatrix} \tag{38}$$

and the orthogonal transformation is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = A \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}$$

These formulæ written in full are known as Rodrigues' equations.[1] They were also known to Euler (1770). Their chief interest is that they give a *rational* solution of the problem, the simplest case, when $n = 2$, being familiar in the form of finding rational lengths for the sides of a right-angle triangle.

To obtain, by any other means, a set of rational values of direction cosines of three mutually perpendicular lines in space referred to rectangular Cartesian axes is a difficult problem, as an attempt will readily show.

EXAMPLES

1. Verify that $A = \frac{1}{3}\begin{bmatrix} 1 & -2 & 2 \\ -2 & 1 & 2 \\ -2 & -2 & -1 \end{bmatrix}$ gives an orthogonal transformation.

2. Prove that if A is the Cayley orthogonal matrix, then $|A| = 1$.

3. If p rows or p columns of the Cayley matrix are multiplied by -1, the result is an orthogonal matrix whose determinant is ± 1 according as p is even or odd.

[Apply the detailed test as in (37)].

4. If J denotes the unit matrix with p negative and $n - p$ positive signs attached to diagonal elements, then $J(I + S)/(I - S)$ is the general orthogonal matrix, whose determinant is ± 1 according as p is even or odd (cf. p. 348).

---

[1] Rodrigues, *Journ. de Liouville de Math.*, **5**, 404–405; cf. Euler, *Comment. Petropol*, **15** (1770), p. 101.

5. Linear Transformation with Absolute Quadric.

An important corollary follows, which concerns the general quadratic relation

$$\Sigma a_{ij} x_i x_j = \Sigma a_{ij} y_i y_j, \quad . \; . \; . \; . \; . \quad (39)$$

analogous to the simpler case already taken. Can a linear transformation of variables

$$X' = [x_1, x_2, \ldots, x_n], \quad Y' = [y_1, y_2, \ldots, y_n] \qquad (40)$$

from X to Y be found, such that the above quadratic relation is identically satisfied? The answer is given by use of a symmetrical matrix $Q = [a_{ij}] = [a_{ji}]$; for the quadratic itself may be denoted by the matrix product

$$X'QX. \quad . \; . \; . \; . \; . \; . \; . \quad (41)$$

For example,

$$[x_1, \; x_2, \; x_3] \begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$= ax_1^2 + bx_2^2 + cx_3^2 + 2fx_2x_3 + 2gx_3x_1 + 2hx_1x_2. \quad (42)$$

Our condition is now $X'QX = Y'QY$, and this is secured by taking as our linear transformation

$$X = \frac{I + SQ}{I - SQ} Y, \qquad (43)$$

S being an arbitrary skew symmetric matrix. For by transposition

$$X' = Y' \frac{I - QS}{I + QS}, \qquad (44)$$

since $Q' = Q$, $S' = -S$. Hence $X'QX = Y'QY$ provided

$$\frac{I - QS}{I + QS} Q \frac{I + SQ}{I - SQ} = Q. \qquad (45)$$

As in (37) the fractional notation is unambiguous. Multiplying fore and aft by $I + QS$ and $I - SQ$ respectively we have

$$(I - QS)\, Q\, (I + SQ) = (I + QS)\, Q\, (I - SQ), \qquad (46)$$

which is true on expansion without the commutative law of multiplication.

This matrix $\dfrac{I + SQ}{I - SQ}$ is a function of a single argument SQ.

Another matrix which has the same property, but which is not a function of one argument, for the order of its factors is non-commutative, is given by the following theorem.

Hermite's Theorem.—*The matrix* $(Q + S)^{-1}(Q - S)$ *gives rise to a linear transformation which leaves the quadric* $X'QX$ *unchanged.*

Proof.—

Let $R = (Q + S)^{-1}(Q - S)$, so that, by the reversal law, its transposed is

$$R' = (Q + S)(Q - S)^{-1}. \quad . \quad . \quad . \quad . \quad (47)$$

Hence

$$\begin{aligned} R'(Q + S)R &= (Q + S)(Q - S)^{-1}(Q + S)(Q + S)^{-1}(Q - S), \\ &= Q + S \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (48) \end{aligned}$$

after cancelling the third and fourth factors and then the second and fifth. Similarly

$$R'(Q - S)R = Q - S. \quad . \quad . \quad . \quad . \quad . \quad . \quad (49)$$

Adding these results we have

$$R'(Q + S + Q - S)R = 2Q$$

or

$$R'QR = Q. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (50)$$

By subtraction,

$$R'SR = S. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (51)$$

Hence as in (45) the requisite condition is satisfied, so proving the theorem.

Corollary I.—*The same matrix leaves the skew symmetric bilinear form*

$$X'SY = \Sigma\, x_i s_{ij} y_j \qquad (s_{ij} = - s_{ji})$$

unchanged, as is indicated by (51).

Corollary II.—*The pencil of bilinear forms*

$$\lambda \Sigma\, x_i a_{ij} y_j + \mu \Sigma\, x_i s_{ij} y$$

whose matrix is $\lambda Q + \mu S$, *is also left unchanged. For by* (50) *and* (51), *if* λ *and* μ *are scalar,*

$$R'(\lambda Q + \mu S)R = \lambda Q + \mu S.$$

6. Group of the Orthogonal Matrix.

THEOREM.—*The product of two orthogonal* n-*rowed matrices is orthogonal.* In fact, if A and B are each orthogonal n-rowed matrices, then

$$AA' = BB' = A'A = B'B = I.$$

Hence

$$ABB'A' = AIA' = AA' = I.$$

Thus if $C = AB$, then $C' = B'A'$ and $CC' = ABB'A' = I$,

which proves the theorem.

This result is obvious geometrically by interpreting each matrix as a suitable rotation of axes about a fixed origin.

Such a result typifies a property of fundamental importance throughout mathematics, namely the group property. In its general form the group is defined as follows.[1]

Definition of a Group.—*A system consisting of a class of elements* A, B, C, . . . , *and one rule of combination, which will be denoted by* o, *is called a group if the following conditions are satisfied*:

(1) *If* A *and* B *are members of the class, whether distinct or not,* A o B *is also a member of the class.*

(2) *The associative law holds, namely*

$$(A \circ B) \circ C = A \circ (B \circ C).$$

(3) *The class contains a member* I *called the identical element, which is such that every member is unchanged when combined with it: thus*

$$A \circ I = I \circ A = A.$$

(4) *Answering to each member* A *is a member* A^{-1}, *called the inverse of* A, *such that*

$$A \circ (A^{-1}) = (A^{-1}) \circ A = I.$$

Simple examples of groups, which obey these conditions, will at once occur to the reader. Positive and negative integers with zero form a group, if the rule of combination is addition. In this case the inverse of A is $-A$, while $I = 0$. So the class of integers is a group for the operation addition. But not so for subtraction; nor even for multiplication, since condition (4) breaks down.

---

[1] Cf. Bôcher, *Higher Algebra* (New York, 1919), p. 82.

Non-singular matrices of the same order form a group for multiplication, the identical element being I.

The totality of all displacements of a rigid plane lamina in its own plane form a group, if we allow the null displacement—no displacement at all—to act as the identical element of the group.

Again, the totality of all linear transformations, from a set X of n variables to a set Y, form a group, provided the matrix M of the transformation is non-singular. Thus if $X = MY$ is a transformation $X \rightarrow Y$, and $Y = NZ$ is another linear transformation, then

$$X = MY = M(NZ) = (MN)Z.$$

Hence (MN) the product of the matrices M, N determines the linear transformation from X direct to Z (§2, p. 59). We may evidently speak of the transformation M (or N, or MN), meaning that for which M is its coefficient matrix, as in §1. So if $|M|$, $|N|$ are non-zero, the inverse transformations exist, and group condition (4) is satisfied. If in particular $N = M^{-1}$, then

$$X = MM^{-1}Z = IZ,$$

which gives the *identical transformation* of the group, namely

$$x_1 = z_1, \; x_2 = z_2, \; \ldots, \; x_n = z_n.$$

7. Dimensions of the Transformation Group.

Consider the two matrices M and N, each with n^2 elements, not entirely alike, so that $M \neq N$. Then $MY \neq NY$, so that we may properly speak of the transformations

$$X = MY, \qquad X = NY$$

as distinct. For they give different values of the set X answering to one value of the set Y. Thus the transformation M is said to have n^2 dimensions, for it is not specified unless all its n^2 elements are given, whereas these n^2 elements determine it uniquely.

Definition of Subgroup.—*A subgroup of a group is an aggregate of members which themselves form a group, with the same rule of combination.*

For our purpose the following examples are important.

The Projective Subgroup.

All matrices ρM, σM, τM . . . where ρ, σ, τ . . . are non-zero scalar factors are members of a group. Each is obtained from another by scalar multiplication, and all the group conditions are satisfied. Applied to a linear transformation $X \rightarrow Y$ such matrices differ merely by multiplying the co-ordinates $y_1, y_2, \ldots, y_n$ by a constant factor. Now this difference is immaterial for *homogeneous* co-ordinates in geometry: accordingly these transformations are indistinguishable. So if the $n^2 - 1$ ratios $\xi_1 : \xi_2 : \ldots : \omega_n$ of the n^2 elements of M are given, one such transformation is determined. The totality of these transformations forms the *projective group* of $(n-1)$-fold space, and therefore its group dimensions are $n^2 - 1$.

The Affine Group.

The matrix

$$M_1 = \begin{bmatrix} \xi_1 & \ldots & \zeta_1 & \omega_1 \\ & & & \\ \xi_{n-1} & & \zeta_{n-1} & \omega_{n-1} \\ 0 & & 0 & \omega_n \end{bmatrix}, \quad \omega_n \neq 0, \quad \Delta = \begin{vmatrix} \xi_1 & \ldots & \zeta_1 \\ \cdot & \cdot \quad \cdot \quad \cdot & \cdot \\ \xi_{n-1} & \ldots & \zeta_{n-1} \end{vmatrix} \neq 0$$

with $n - 1$ zeros in the nth row defines a group of transformations. For the product of two such still has the requisite zeros; and each group property, including the existence of an identical member, is secured. Here then is a subgroup of the general transformation group. Since M_1 has $n^2 - n + 1$ arbitrary elements, this number gives the dimensions of the affine group for $(n-1)$-fold space.

The Affine Group with a Fixed Point.

This is defined by

$$M_2 = \begin{bmatrix} \xi_1 & \ldots & \zeta_1 & 0 \\ & & & \\ \xi_{n-1} & & \zeta_{n-1} & 0 \\ 0 & \ldots & 0 & \omega_n \end{bmatrix}, \quad \omega_n \neq 0, \quad \Delta \neq 0$$

with $(n-1)^2 + 1$ group dimensions. Here there are $n - 1$

zeros in both the last row and column. Again the group properties hold.

The Orthogonal Affine Group with a Fixed Point.

The matrix M_{12} of type M_2 *when the minor* $\begin{bmatrix} \xi_1 & \dots & \zeta_1 \\ . & . & . \\ \xi_{n-1} & \dots & \zeta_{n-1} \end{bmatrix}$ *is orthogonal* defines a group, as is easily verified. The $\frac{1}{2}n(n-1)$ conditions needed to make this minor orthogonal cut down the group dimensions to

$$(n-1)^2+1-\tfrac{1}{2}n(n-1)=\tfrac{1}{2}(n^2-3n+4).$$

EXAMPLES

1. The affine group leaves the equation of a certain prime (linear form) absolute; namely, $x_n = 0$ becomes $x_n' = 0$ when $x \longrightarrow x'$.

If $n = 3$, this is illustrated by taking x_1/x_3, x_2/x_3 as Cartesian oblique co-ordinates and regarding the transformation as a change of axes.

2. Keeping the same axes and regarding the transformation as a change of figure, the point x moving to x', prove that the affine group ($n = 3$) changes point to point, line to line, and parallel lines to parallel lines.

3. The M_2 group can be regarded as a change of axes without shifting the origin.

8. Induced Compound Transformations.

In §8, p. 86, certain compound co-ordinate sets $\pi_2, \pi_3, \dots$, $p_{n-1}, p_{n-2}, \dots$ were introduced as a direct application of the determinant theory. We now proceed to develop their properties in relation to the Sylvester theory of cogredience.

Let $x, y, z, \dots, s, t$ be any number k of cogredient variables, r of which we choose to form a matrix of r rows and n columns, say

$$[\rho_r] = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \end{bmatrix}.$$

This has $\binom{n}{r}$ determinants of order r, making a set which we call ρ_r, provided $r < n$.

Definition.—*The* r*-rowed determinants of the set* ρ_r *are called* r*th compound point co-ordinates.*

Accordingly the sets π_2, π_3 . . . are particular second, third, . . . compound point co-ordinates. If we abbreviate the set $\pi_2 = (xy)_{\alpha\beta}$ as $\overline{xy}$, and π_3 as $\overline{xyz}$, and so on, we consider all the various sets

$$\overline{xy},\quad \overline{xz},\quad \dots,\quad \overline{xt},\quad \overline{yz},\quad \dots,\quad \overline{st}$$

as second compounds: and similar remarks apply to rth compounds. Thus from k given points $x, y, \dots, t$ we derive $\binom{k}{r}$ rth compounds.

In particular, if $r = n - 1$ the rth compound π_{n-1} is a prime (§8, p. 86). This is true of any $(n-1)$th compound. Also if $r = 1$ we revert to the original point type x or y . . . or t. And once more, if $r = n$, the matrix $[\rho_r]$ is square and has a single determinant $|\rho_n|$. This gives a set of n points which form a simplex provided $|\rho_n| \neq 0$: otherwise the points are linearly related.

We now come to an important theorem.

THEOREM.—*A linear transformation* T *of cogredient variables* x, y, . . . *to* x′, y′, . . . *respectively, induces a linear transformation upon all their compounds* $\overline{\mathrm{xy}}$, $\overline{\mathrm{xyz}}$, . . . , *such that all* r*th compounds are cogredient.*

Proof.—

This follows immediately from the theorem of corresponding matrices (§4, p. 79). For by (4), p. 148, which we write shortly as $x_1' = \xi_x$, $x_2' = \eta_x$, &c., we have for cogredient variables

$$y_1' = \xi_y,\quad z_1' = \xi_z, \dots,\quad y_n' = \omega_y,\quad z_n' = \omega_z, \dots$$

Hence

$$(x'y')_{12} = \begin{vmatrix} \xi_x & \xi_y \\ \eta_x & \eta_y \end{vmatrix} = (\xi\eta \mid xy) = \Sigma\,(\xi\eta)^{ij}(xy)_{ij},$$

which shows that the new compound co-ordinate $(x'y')_{12}$ is a linear function of all the old $(xy)_{ij}$. Similarly, if θ, ϕ belong to the αth and βth rows of (4),

$$(x'y')_{\alpha\beta} = (\theta\phi \mid xy) = \Sigma\,(\theta\phi)^{ij}(xy)_{ij}.$$

Hence the second compound $\overline{x'y'}$ is transformed linearly to $\overline{xy}$ by a matrix whose elements are the second compounds of the elements of M^{-1} which transforms x' to x and y' to y. The same matrix arises whatever pair among x, y, z, . . . , t is first selected. So all second compounds are cogredient.

Likewise for third compounds

$$(x'y'z')_{123} = (\xi\eta\zeta \mid xyz) = \Sigma(\xi\eta\zeta)^{ijk}(xyz)_{ijk}$$

leading to a similar result; and so on until the $(n - 1)$th compound is reached, in which case the typical equation is

$$(x'y'z' \ldots s')_{23\ldots n} = \Sigma(\eta\zeta \ldots \omega)^{ij\ldots l}(xyz \ldots s)_{ij\ldots l}.$$

The coefficients in this series are proportional to

$$\xi_1, \xi_2, \ldots, \xi_n,$$

which leads back to the result that the transformation of $(n-1)$th compounds is cogredient with that of u and contragredient to x.

This puts the Sylvester-Cauchy theorem (§9, p. 87) on compound determinants in a new light. For we have now arrived at a system of compound matrices, say M, M_2, M_3, . . . , M_{n-1}, whose determinants $|M|$, $|M_2|$, . . . are what have been called compound determinants. Since each determinant, according to this theorem, is a power of $|M|$, it follows that none of these compound linear transformations are singular unless that of x itself is.

Corollary I.—*The correlative compounds* p_r *undergo linear transformation, such that* p_r *and* π_r *are contragredient.*

This follows at once from §8, p. 86.

Corollary II.—*If the transformation* $\mathrm{T} : \mathrm{x} \rightarrow \mathrm{x}'$ *is orthogonal, so also is each compound transformation.*

For if α, β, γ . . . denote any of the rows ξ, η, . . . , ω in the transformation matrix M, then

$$(\alpha \mid \alpha) = 1, \quad (\alpha \mid \beta) = 0 \quad \alpha \neq \beta,$$

when M is orthogonal: whence, by the theorem of corresponding matrices with r letters both before and after the vertical line,

$$(\alpha\beta \ldots \mid \gamma\delta \ldots) = 1 \text{ or } 0$$

according as $\alpha = \gamma$, $\beta = \delta$, . . . or at least one of γ, δ, . . . differs from α, β, ... These conditions at once imply that the rth compound M_r is orthogonal.

9. Connexion between Matrices and Quaternions.

The theory of four-rowed orthogonal matrices is intimately connected with that of quaternions. If we introduce into non-commutative algebra three elements i, j, k called complex units, defined solely by the equations

$$\begin{gathered} i^2 = j^2 = k^2 = -1 \\ jk = i = -kj, \quad ki = j = -ik, \quad ij = k = -ji, \end{gathered} \tag{52}$$

then a quaternion is a linear function of i, j, k

$$q = ix + jy + kz + t,$$

where x, y, z, t are scalar.

If x, y, z, t belong to the field of real numbers, q is called a real quaternion; if to the field of complex numbers $(\alpha + \sqrt{-1}\,\beta)$, q is a complex quaternion.

The quaternion

$$q' = -ix - jy - kz + t$$

is called the *conjugate* of q; it satisfies the scalar condition

$$qq' = q'q = x^2 + y^2 + z^2 + t^2$$

analogous to the property of conjugate complex numbers $\alpha + \sqrt{-1}\,\beta$ and $\alpha - \sqrt{-1}\,\beta$. This quadratic expression $x^2 + y^2 + z^2 + t^2$ is called the *norm* of q.

EXAMPLES

1. Prove that the two-row matrices

$$I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad i = \begin{bmatrix} \sqrt{-1} & 0 \\ 0 & -\sqrt{-1} \end{bmatrix}, \quad j = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad k = \begin{bmatrix} 0 & \sqrt{-1} \\ \sqrt{-1} & 0 \end{bmatrix}$$

satisfy the above properties and verify that i, j, k as defined by (52) also satisfy the associative laws.

2. A quaternion q is expressible as a two-rowed matrix

$$q = \begin{bmatrix} t + \varepsilon x, & y + \varepsilon z \\ -y + \varepsilon z, & t - \varepsilon x \end{bmatrix}$$

where ε denotes $\sqrt{-1}$.

3. Prove that the matrix product qq' is scalar, where q' is the conjugate of q.

4. If z, $\bar{z}$ are conjugate complex numbers, and also w, $\bar{w}$, then $\begin{bmatrix} z, & w \\ -\bar{w}, & \bar{z} \end{bmatrix}$ is a quaternion.

Prove the reversal law for the conjugate of a product of quaternions p, q, namely

$$(pq)' = q'p'.$$

5. *The norm of a product is the product of the norms.* This generalizes the well-known theorem for the product of moduli of complex numbers.

Justify the steps in the following proof:

If $r = pq$, then $r' = q'p'$. So

$$rr' = pqq'p' = p(qq')p' = (pp')(qq').$$

6. Taking $p = i\alpha + j\beta + k\gamma + \delta$, $q = ix + jy + kz + t$ express $r = pq$ in full as a quaternion. Hence by **5**, prove the identity

$$(\alpha^2 + \beta^2 + \gamma^2 + \delta^2)(x^2 + y^2 + z^2 + t^2) = X^2 + Y^2 + Z^2 + T^2$$

where

$$\begin{aligned} X &= \delta x - \gamma y + \beta z + \alpha t \\ Y &= \gamma x + \delta y - \alpha z + \beta t \\ Z &= -\beta x + \alpha y + \delta z + \gamma t \\ T &= -\alpha x - \beta y - \gamma z + \delta t. \end{aligned}$$

7. If $x^2 + y^2 + z^2 + t^2 = 1 = \alpha^2 + \beta^2 + \gamma^2 + \delta^2$, prove that

$$A = \begin{bmatrix} \delta & -\gamma & \beta & \alpha \\ \gamma & \delta & -\alpha & \beta \\ -\beta & \alpha & \delta & \gamma \\ -\alpha & -\beta & -\gamma & \delta \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} t & z & -y & x \\ -z & t & x & y \\ y & -x & t & z \\ -x & -y & -z & t \end{bmatrix}$$

are orthogonal matrices.

8. If $\begin{bmatrix} X_1 & X_2 & & X_4 \\ Y_1 & Y_2 & & Y_4 \\ Z_1 & Z_2 & Z_3 & \\ T_1 & T_2 & T_3 & T_4 \end{bmatrix}$ denotes the product AB, prove that the sum of the squares of elements in each column is

$$(\alpha^2 + \beta^2 + \gamma^2 + \delta^2)(x^2 + y^2 + z^2 + t^2).$$

CHAPTER X

GENERAL PROPERTIES OF INVARIANTS

1. Linear Transformation of the General Form of Order p.

Certain theorems apply equally well to forms in n variables $x_1, x_2, \ldots, x_n$ as to binary forms; so we shall now consider them.

Let
$$f = \sum_{i=1} c_i x_1^{\alpha_i} x_2^{\beta_i} \ldots x_n^{\nu_i} \quad . \quad . \quad . \quad . \quad (1)$$

be a form of order p, so that

$$\alpha_i + \beta_i + \ldots + \nu_i = p \quad . \quad . \quad . \quad . \quad (2)$$

for each term of f. We replace c_i by a multinomial coefficient together with an arbitrary coefficient a_i, such that

$$c_i = \frac{p!}{\alpha_i!\ \beta_i! \ldots \nu_i!} a_i. \qquad (3)$$

Let N be the number of different terms in the general p-ic f, that is the number of different values of the matrix

$$[\alpha_i, \beta_i, \ldots, \nu_i]. \qquad (4)$$

Then N is also the number of different terms in the special form when $a_1 = a_2 = \ldots = a_N = 1$, namely

$$(x_1 + x_2 + \ldots + x_n)^p. \quad . \quad . \quad . \quad . \quad (5)$$

With this understanding we write

$$f = (a_1, a_2, \ldots, a_N)(x_1, x_2, \ldots x_n)^p = f(x) \qquad (6)$$

in the contracted functional notation.

Now let $T : x \longrightarrow x'$ be a linear transformation with n equations

$$x_i = \xi_i x_1' + \eta_i x_2' + \ldots + \omega_i x_n', \quad . \quad . \quad (7)$$

where the square matrix M of the n^2 coefficients $\xi_1, \ldots, \omega_n$ has a non-zero determinant $|M|$. Subject to this sole condition, the coefficients are arbitrary independent real or complex numbers.

As in binary forms the effect of this transformation upon the p-ic $f(x)$ is to produce a new p-ic $f'(x')$. Thus

$$\begin{aligned} f = f(x) = f'(x') &= (a_1, a_2, \ldots, a_N \between x_1, x_2, \ldots, x_n)^p \\ &= (a_1', a_2', \ldots, a_N' \between x_1', x_2', \ldots, x_n')^p, \quad . \quad (8) \end{aligned}$$

defining a set of N new coefficients $[a']$ analogous to (3). In fact a_i' is the coefficient of $x_1'^{\alpha_i} x_2'^{\beta_i} \ldots x_n'^{\nu_i}$ after removing the multinomial factor. But this is found on the left-hand side by picking out the required terms in the expansion of each separate term. For our present purpose it is sufficient to observe that *each* a_i' *is a linear function of* $a_1, a_2, \ldots, a_N$. We typify this by

$$T : x \rightarrow x', \qquad T_a : a \rightarrow a'.$$

2. Projective Invariants.

Definition of Invariant.—*A polynomial function* I(a, b, . . .) *of the coefficients* a, b, . . . *of forms* f(x), g(x), . . . *is a polynomial projective invariant if*

$$I(a', b', \ldots) = \phi(\xi) I(a, b, \ldots) \quad . \quad . \quad . \quad (9)$$

identically, where $\phi(\xi)$ *is a factor depending solely on the* n^2 *coefficients* $\xi_1, \ldots, \omega_n$ *of the transformation* $x \rightarrow x'$.

Such a function is a *relative projective invariant*, to give it its full accepted title, but briefly it is called an invariant. We prove a few theorems which hold of such functions I.

Theorem I.—*The factor* $\phi(\xi)$ *is a positive integral power of the modulus* $|M|$ *of the transformation* $x \rightarrow x'$; *namely*

$$\phi(\xi_1, \ldots, \omega_n) = |M|^w = (\xi\eta \ldots \omega)^w. \quad . \quad . \quad (10)$$

Proof.—

For if $I(a)$ denote such an invariant of $f(x)$ then

$$I(a') = \phi(\xi) I(a).$$

But suppose we start with $f'(x')$, then $I(a')$ is an invariant of

$f'(x')$. If we now transform f' back to f by the inverse transformation $x' \to x$ (cf. (5), p. 69),

$$x_1' = \frac{X_1}{|M|}x_1 + \frac{X_2}{|M|}x_2 + \ldots + \frac{X_n}{|M|}x_n$$

$$x_2' = \frac{Y_1}{|M|}x_1 + \frac{Y_2}{|M|}x_2 + \ldots + \frac{Y_n}{|M|}x_n$$

and the condition analogous to (9) is

$$I(a) = \phi\left(\frac{X}{|M|}\right) I(a').$$

Hence if $I(a) \neq 0$, we obtain by multiplying these results

$$\phi(\xi)\,\phi\left(\frac{X}{|M|}\right) = 1 \quad . \quad . \quad . \quad . \quad . \quad (11)$$

Since $I(a')$ is rational and integral, so also must $\phi\left(\frac{X}{|M|}\right)$ be.

Hence we can clear the denominator of (11) on multiplying through by a suitable power $|M|^s$ and, after expressing each co-factor X of $|M|$ in terms of the elements $\xi_1 \ldots$, we obtain

$$\phi(\xi)\,\psi(\xi) = |M|^s, \quad . \quad . \quad . \quad . \quad . \quad (12)$$

where both ϕ and ψ are polynomials in their arguments. But $|M|$ is an arbitrary determinant and therefore has no factors rational and integral in its elements. Consequently both $\phi(\xi)$ and $\psi(\xi)$ are powers of $|M|$. Thus $\phi(\xi) = |M|^w = \Delta^w$.

This index w is called the *weight* of the invariant I.

Corollary.—*If* I_1, I_2 *are two invariants of weight* w_1, w_2 *respectively, then*

$$I_1' = \Delta^{w_1} I_1, \qquad I_2' = \Delta^{w_2} I_2.$$

Hence the product $I_1 I_2$ *is an invariant of weight* $w_1 + w_2$. *The quotient* I_1/I_2 *satisfies the condition of invariancy, and is called a rational invariant of weight* $w_1 - w_2$.

An *algebraic invariant* is the root of an equation

$$(I_0, I_1, \ldots, I_p \between z, 1)^p = 0,$$

where each coefficient is a rational integral invariant.

The sum of two invariants I_1, I_2 is only invariant if their weights are equal. For

$$I_1' + I_2' = \Delta^{w_1} I_1 + \Delta^{w_2} I_2,$$

and if the left-hand side is invariant it has just been proved equal to $\Delta^w(I_1 + I_2)$. Hence $w = w_1 = w_2$.

We sum this up by saying an invariant is isobaric.

EXAMPLE

Show that the present definition of weight w agrees in the binary case with the definition already introduced in §4, p. 134.

3. Homogeneity of Invariants.

Consider a number of given ground forms $f(x)$, $g(x)$, ... whose coefficient sets are $[a]$, $[b]$, ... Let I be a simultaneous invariant of weight w, so that

$$I(a', b', \ldots) = \Delta^w I(a, b, \ldots).$$

We shall prove that it may be sorted out uniquely into a number of terms homogeneous in each set $[a]$, $[b]$, ..., each such term being an invariant.

Theorem II.—*Every simultaneous invariant can be expressed in one and only one way as a sum*

$$I = I' + I'' + \ldots I^{(\sigma)}$$

of invariants $\mathrm{I}^{(h)}$ *which are each of weight* w *and homogeneous in each set of coefficients involved.*

Proof.—

First let all terms be reduced as far as possible, terms with the same index set, §1 (4), being collected into one term. If I is not homogeneous in each set $[a]$, $[b]$, ... let it be written

$$I = I_1(a) + I_2(a) + \ldots + I_s(a),$$

where each term in this sum is homogeneous in each set.

Now by definition we have

$$I(a') = \phi \times \{I_1(a) + I_2(a) + \ldots + I_s(a)\}.$$

Therefore

$$I_1(a') + I_2(a') + \ldots + I_s(a') = \phi \times \{I_1(a) + \ldots I_s(a)\}$$

identically. Also ϕ is independent of $a, b, \ldots$, while a' is linear in a; b' in b; .... Hence the only part on the left-hand side which is of the same degree in a as $I_1(a)$ on the right is $I_1(a')$; so that

$$I_1(a') = \phi I_1(a).$$

Thus $I_1(a)$ is an invariant.

For example

$$a_0a_2 - a_1^2 + a_0b_2 - 2a_1b_1 + b_0a_2$$

is a simultaneous invariant of the two binary quadratics

$$(a_0, a_1, a_2 \between x_1, x_2)^2 \quad \text{and} \quad (b_0, b_1, b_2 \between x_1, x_2)^2,$$

but it is the sum of the two expressions

$$a_0a_2 - a_1^2 \quad \text{and} \quad a_0b_2 - 2a_1b_1 + a_2b_0,$$

each of which is homogeneous in the two sets of coefficients. Calling these invariants I_1 and I_2, they both have weight two, but they differ in degree. The *degree* of an invariant of a single form is its degree in the coefficients of the form. So I_1 has degree two and weight two. In keeping with this definition, I_2 is said to have partial degrees (1, 1) in the respective sets of coefficients, and again its weight is two.

4. Ground Forms.

It is natural to extend the definition of a ground form (p. 138) to include the case of several such forms as in the above example. The coefficients of terms in the forms are *ground coefficients*.

It should now be clear that three essential things are involved in the invariant theory: the ground form, the transformation, and the invariant. In its general aspect the problem before us is to discover whether a function, say $I(a)$, exists, and if so how many such functions exist. To these questions a general answer can be given, not unlike the corresponding answer to the question whether a given equation, algebraic or differential, has a solution. The results are crystallized in the great theorems which follow later, associated with the names of Clebsch, Gordan, and Hilbert.

5. Symbolic Notation.

The reader is already familiar with differential operators which combine very like ordinary numbers. Suppose, for example, that x_1, x_2 are independent of x and y. Then we may write

$$\left(x_1\frac{\partial}{\partial x}+x_2\frac{\partial}{\partial y}\right)^p f(x, y)$$
$$=x_1{}^p\frac{\partial^p f}{\partial x^p}+px_1{}^{p-1}x_2\frac{\partial^p f}{\partial x^{p-1}\partial y}+\ldots+x_2{}^p\frac{\partial^p f}{\partial y^p} \qquad (13)$$

identically, provided p is a positive integer and f a function capable of such successive differentiation.

In particular let

$$f=a_0x^p+pa_1x^{p-1}y+\ldots+a_py^p, \quad . \quad . \quad (14)$$

then each pth derivate of f is a single term, and, in fact,

$$\frac{\partial^p f}{\partial x^{p-r}\partial y^r}=p!\,a_r, \qquad r=0, 1, 2, \ldots. p. \quad . \quad (15)$$

Hence identity (13) now takes the form

$$\left(x_1\frac{\partial}{\partial x}+x_2\frac{\partial}{\partial y}\right)^p f(x, y)$$
$$=p!(a_0x_1{}^p+pa_1x_1{}^{p-1}x_2+\ldots+a_px_2{}^p). \qquad (16)$$

Let us now introduce the following notation:

$$\alpha_1{}^{p-r}\alpha_2{}^r=\frac{\partial^p f}{\partial x^{p-r}\partial y^r}\div p!, \quad r=0, 1, 2, \ldots, p, \qquad (17)$$

so that for extreme values of r,

$$\alpha_1{}^p=\frac{\partial^p f}{\partial x^p}\div p!, \quad \alpha_2{}^p=\frac{\partial^p f}{\partial y^p}\div p!.$$

Then owing to the convenient fact that the differential operators $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ combine as numbers and obey the index law, so also do α_1, α_2. This becomes plainer if we take an actual example, say

$$\frac{\partial^5 f}{\partial x\,\partial y\,\partial x^2\,\partial y}\div 5!,$$

which would occur if $p = 5$. This would be written as

$$\alpha_1 \alpha_2 \alpha_1^2 \alpha_2.$$

Manifestly we should have such relations as

$$\alpha_1 \alpha_2 \alpha_1^2 \alpha_2 = \alpha_1^3 \alpha_2^2 = \alpha_2^2 \alpha_1^3 = \&c.;$$

for they are all ways of writing the same quantity

$$\frac{\partial^5 f}{\partial x^3 \partial y^2} \cdot \quad 5!.$$

So we have introduced two symbols α_1, α_2 which have the properties of ordinary numbers with this proviso: they only occur in a product of degree p, involving $p - r$ factors α_1 and r factors α_2: otherwise they are undefined.

Now let us extend this definition. Let α_1 and α_2 occur in a product of degree 1, 2, 3, . . . , or p, and behave like ordinary numbers, but let a product of degree greater than p be undefined and therefore meaningless. There is clearly no contradiction involved in such a restriction.

If we substitute from (17) in (15) we obtain the elegant result

$$\alpha_1^p = a_0, \ \alpha_1^{p-1} \alpha_2 = a_1, \ \ldots, \ \alpha_1^{p-r} \alpha_2^r = a_r, \ \ldots, \ \alpha_2^p = a_p. \quad (18)$$

The result when put in (16) gives

$$(x_1 \alpha_1 + x_2 \alpha_2)^p = a_0 x_1^p + p a_1 x_1^{p-1} x_2 + \ldots + a_p x_2^p. \quad (19)$$

Since x_1, x_2 are independent of x and y, they combine with the differential operators and therefore with α_1 and α_2 as with ordinary numbers. So we may write

$$(\alpha_1 x_1 + \alpha_2 x_2)^p = a_0 x_1^p + p a_1 x_1^{p-1} x_2 + \ldots + a_p x_2^p. \quad (20)$$

This is an identity for x_1, x_2, obviously agreeing with relations (18).

These α_1, α_2 are the Clebsch-Aronhold symbols, founded on the hyperdeterminants of Cayley, which have proved to be of the utmost value in developing the general theorems of the invariant theory. Now that they have been defined we can dispense with all that precedes (18) and (20) by making the following doctrine of these symbols:

The symbols α_1, α_2 behave as ordinary numbers. They have *no actual meaning as numbers except when they occur in a product involving exactly* p *of them.*

Thus the symbols express the coefficients $a_0, a_1, \ldots, a_p$ of the binary p-ic f explicitly and uniquely, and any linear function of the coefficients can be written unambiguously by means of the symbols. Indeed *they express the binary* p*-ic itself as a perfect* p*th power of a symbolic linear form* $\alpha_1 x_1 + \alpha_2 x_2$.

If in particular the binary p-ic happens to be a perfect pth power, the symbols represent actual numbers. This is called the *scalar instance* of the general symbolic form.

6. Symbols for Forms in Three or More Variables.

Exactly the same methods may be used to denote a homogeneous form in three or more variables, $x_1, x_2, x_3, \ldots$ by means of symbols $\alpha_1, \alpha_2, \alpha_3, \ldots$.

From an identity analogous to (20), we should arrive at the result

$$(\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3)^p = \Sigma \frac{p!}{i!j!k!} a_{ijk} x_1^i x_2^j x_3^k, \quad . \quad (21)$$

where $i+j+k=p$, the summation extending to all different values of the index matrix $[i, j, k]$. The only essential difference here is in choosing a suitable notation for the coefficient of the ternary p-ic.

$$f = a_{p00} x^p + \ldots + a_{00p} z^p,$$

which now takes the place of (14). Whatever principle of suffix or other notation is adopted on the right-hand side of (21), it is agreed that a product of p symbols

$$\alpha_1^i \alpha_2^j \alpha_3^k,$$

multiplied by the trinomial coefficient $p!/i!j!k!$ actually represents the coefficient of $x_1^i x_2^j x_3^k$ in this ternary p-ic.

If $p = 2$, the coefficients are best denoted by double suffixes in all cases involving many variables. Thus

$$f = \sum_i \sum_j a_{ij} x_i x_j$$

$$i = 1, 2, \ldots, n, \quad j = 1, 2, \ldots, n \quad a_{ij} = a$$

is the quadratic form in n homogeneous variables.

For instance, in this notation the areal or homogeneous equation of a conic is

$$f = a_{11}x_1^2 + a_{22}x_2^2 + a_{33}x_3^2 + 2a_{23}x_2x_3 + 2a_{31}x_3x_1 + 2a_{12}x_1x_2,$$

which is symbolically written

$$(a_1x_1 + a_2x_2 + a_3x_3)^2.$$

This leads to the very simple definition of symbols for a quadratic, namely

$$a_ia_j = a_{ij}.$$

So also, in accordance with the defined behaviour of the symbols,

$$a_ja_i = a_{ji} = a_{ij}.$$

Similarly the quaternary quadratic is denoted by

$$(a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4)^2,$$

and the general quadratic by

$$(a_1x_1 + a_2x_2 + \ldots + a_nx_n)^2,$$

where as before

$$a_ia_j = a_{ij}.$$

Cubic Forms.—Next if $p = 3$, a triple suffix notation is convenient, so that the general cubic is symbolized by

$$(a_1x_1 + a_2x_2 + \ldots + a_nx_n)^3,$$

where $a_ia_ja_k$ gives the coefficient of $x_ix_jx_k$, apart from the multinomial coefficient, in this case 1, 3, or 6 according as $i = j = k$, or two only are equal, or all differ.

The General Form of Order p.—This is best denoted by attaching a group of p suffixes for a coefficient. Thus

$$f = \sum_i \ldots \sum_m a_{ijk\ldots m}\; x_ix_jx_k \ldots x_m,$$

the summation extending from 1 to n for each of the p suffixes $i, j, k, \ldots, m$. This will give all possible products $x_i \ldots x_m$ of degree p in the set $\{x_1, \ldots, x_n\}$: also each term obtained by deranging the suffixes of a given term will be of the same kind. We may therefore define all different permutations of suffixes in $a_{ijk\ldots m}$ to be equal. This allows, as a simple consequence, the symbolic form

$$(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_n x_n)^p$$

to represent f; so that, by equating coefficients,

$$\alpha_i \alpha_j \ldots \alpha_m = a_{ijk\ldots m} = a_{jik\ldots m} = \&c.$$

And lastly, if we carefully distinguish between single suffixes and multiple suffixes, we may with advantage use the letter a rather than α for the symbol, and write

$$a_i a_j \ldots a_m = a_{ijk\ldots m}$$

for the typical coefficient of the p-ic f, now symbolized by

$$(a_1 x_1 + a_2 x_2 + \ldots + a_n x_n)^p.$$

Here the *actual* coefficient $a_{ijk\ldots m}$ is said to be resolved into its *symbolic* factors $a_i, a_j, \ldots, a_m$: and conversely, a symbolic product is only an actual coefficient if it contain exactly p factors $a_i, a_j, \ldots$. In particular if $p = 1$, or if f itself is a perfect pth power, this distinction breaks down, and the form provides its own symbol. Such is called by Professor Bell the *scalar instance* of the symbolic expression.

For binary forms, this notation is unsuitable, since single suffixes do duty for the actual coefficients.

7. Polar Forms.

First we contract the notation $\alpha_1 x_1 + \alpha_2 x_2$ of the binary case to α_x, so that the binary p-ic is denoted by α_x^p. Thus

$$\left.\begin{aligned}\alpha_x^p &= (a_0, a_1, \ldots, a_p \between x_1, x_2)^p \\ &= a_0 x_1^p + p a_1 x_1^{p-1} x_2 + \ldots + a_p x_2^p\end{aligned}\right\}. \quad . \quad . \quad . \quad . \quad (22)$$

Then an excellent example of the use of these symbols is in the polar process.

Since α_1, α_2 behave as ordinary numbers, we have

$$\frac{\partial}{\partial x_1}\alpha_x^p = p\alpha_x^{p-1}\alpha_1, \quad \frac{\partial}{\partial x_2}\alpha_x^p = p\alpha_x^{p-1}\alpha_2. \quad . \quad (23)$$

Multiply by y_1, y_2 respectively; add and then divide by p. Hence

$$\frac{1}{p}\left(y\frac{\partial}{\partial x}\right)a_x^p=\frac{1}{p}\left(y_1\frac{\partial}{\partial x_1}+y_2\frac{\partial}{\partial x_2}\right)a_x^p=a_x^{p-1}a_y. \quad (24)$$

This is the first polar of a_x^p with regard to y. Similarly the second polar is

$$\frac{1}{p(p-1)}\left(y\frac{\partial}{\partial x}\right)^2 a_x^p=a_x^{p-2}a_y^2, \quad . \quad . \quad (25)$$

and so on. The rth polar is

$$a_x^{p-r}a_y^r. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (26)$$

The process ends in p steps, for then all higher differentiation produces zero.

Exactly the same notation $a_x^{p-r}a_y^r$ *denotes the* r*th polar of the* p*-ic in* n *homogeneous variables.*

This may at once be verified.

Further, we may have mixed polar forms involving anything up to p different sets $[x]$, $[y]$, $[z]$, ...

EXAMPLES

1. The quadratic $\alpha_x^2=a_0x_1^2+2a_1x_1x_2+a_2x_2^2$ has a polar

$$\alpha_x\alpha_y=a_0x_1y_1+a_1(x_1y_2+x_2y_1)+a_2x_2y_2.$$

2. The cubic $\alpha_x^3=(a_0, a_1, a_2, a_3 \between x_1, x_2)^3$ has two intermediate polars $\alpha_x^2\alpha_y$, $\alpha_x\alpha_y^2$, and one mixed polar $\alpha_x\alpha_y\alpha_z$. Find their expressions in full.

3. The ternary cubic in canonical form is $x_1^3+x_2^3+x_3^3+6mx_1x_2x_3$. What are its polars of type $\alpha_x^2\alpha_y$, $\alpha_x\alpha_y\alpha_z$?

Ans. $x_1^2y_1+x_2^2y_2+x_3^2y_3+2mx_2x_3y_1+2mx_3x_1y_2+2mx_1x_2y_3$.
$x_1y_1z_1+x_2y_2z_2+x_3y_3z_3+m(x_1y_2z_3+x_1y_3z_2+x_2y_1z_3+x_2y_3z_1$
$+x_3y_1z_2+x_3y_2z_1)$.

4. If α_x^2 denote $(\alpha_1x+\alpha_2y+\alpha_3)^2$, in order to symbolize the conic $ax^2+2hxy+by^2+2gx+2fy+c=0$, what does $\alpha_x\alpha_{x'}=0$ symbolize?

Ans. $(ax+hy+g)x'+(hx+by+f)y'+gx+fy+c=0$.

5. Prove symbolically that if the polar of point P for a conic passes through Q, that of Q passes through P.

6. Prove symbolically that when the sets $[x]$, $[y]$, . . . are all the same as the original $[x]$, each polar reverts to the original form.

Ans. $\alpha_x\alpha_y\alpha_z\ldots=\alpha_x\alpha_x\alpha_x\ldots=\alpha_x^p$.

7. If $(\alpha_1x+\alpha_2y+\alpha_3)^p=0$ symbolizes a plane curve of order p in Cartesian co-ordinates, and $\alpha_x^{p-r}\alpha_{x'}^r=0$ is defined as its rth polar for the point (x', y'), prove that the rth polar is a curve of order $p - r$, which only passes through (x', y') if (x', y') is on the original curve.

8. **Equivalent Symbols.**

One objection to these Clebsch-Aronhold symbols will probably occur to the reader. What are we to make of a form of degree two, or more, in the coefficients of a given ground form f expressed symbolically as α_x^p? For by definition we have a set of n symbols $\alpha_1, \alpha_2, \ldots, \alpha_n$ which can only be utilized to express an actual coefficient in products p at a time. Thus a polynomial of degree p in the symbols α is equivalent to a form linear in the actual coefficients of the ground form f. The simplest way of meeting this difficulty is first to consider simultaneous sets of coefficients, and functions which are linear in each of these sets. Suppose, for instance, we have two binary quadratics

$$\left.\begin{aligned} U &= a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2 \\ V &= b_0 x_1^2 + 2b_1 x_1 x_2 + b_2 x_2^2 \end{aligned}\right\}, \quad . \quad . \quad . \quad (27)$$

written symbolically as α_x^2 and β_x^2 respectively, where $\beta_x^2 = (\beta_1 x_1 + \beta_2 x_2)^2$, and β_1, β_2 are symbols referring exclusively to the coefficients b_0, b_1, b_2; in particular

$$\beta_1^2 = b_0, \quad \beta_1\beta_2 = b_1, \quad \beta_2^2 = b_2. \qquad (28)$$

Then manifestly any expression bilinear in the a's and b's can be expressed by a suitable combination of α's and β's. Thus $a_0 b_2 = \alpha_1^2 \beta_2^2$, or, again,

$$a_0 b_2 + a_2 b_0 - 2a_1 b_1 = (\alpha_1\beta_2 - \alpha_2\beta_1)^2 = (\alpha\beta)^2, \text{ say.} \quad (29)$$

Conversely $(\alpha\beta)^2$, being quadratic in both α_1, α_2 and β_1, β_2 can be expressed unambiguously in terms of the coefficients a_0, a_1, a_2 and b_0, b_1, b_2.

But suppose that the given quadratics U, V are identical. Then

$$a_0 = b_0, \quad a_1 = b_1, \quad a_2 = b_2,$$

while

$$a_0 b_2 + a_2 b_0 - 2a_1 b_1 = 2(a_0 a_2 - a_1^2).$$

This last expression, D_{11} say, which is the discriminant of the quadratic U, appears here as a particular case of a bilinear invariant D_{12}, where

$$D_{12} = a_0 b_2 + a_2 b_0 - 2a_1 b_1 .$$

This gives us a clue for the symbolic expression of D_{11}: *we express* $\mathrm{a}_0\mathrm{a}_2 - \mathrm{a}_1^2$ *symbolically by using the Aronhold operator*

$$\left(b\frac{\partial}{\partial a}\right) \equiv b_0\frac{\partial}{\partial a_0} + b_1\frac{\partial}{\partial a_1} + b_2\frac{\partial}{\partial a_2}$$

on D_{11} *to render it linear in both sets* [a] *and* [b], *after which we are at liberty to use symbols* α *and* β. There are now two α's and two β's in every term of the result $(\alpha\beta)^2$, and the letters α and β are said to be *equivalent symbols.*

We therefore write

$$\begin{aligned} U &= \alpha_x^2 &&= \beta_x^2, \\ a_0 &= \alpha_1^2 &&= \beta_1^2, \\ a_1 &= \alpha_1\alpha_2 &&= \beta_1\beta_2, \\ a_2 &= \alpha_2^2 &&= \beta_2^2. \end{aligned}$$

Similarly for the binary p-ic

$$\begin{aligned} f &= \alpha_x^p = \beta_x^p \\ &= (a_0, a_1, a_2, \ldots, a_p \between x_1, x_2)^p. \end{aligned}$$

Any product of degree two in the coefficients, say $a_i a_j$, is symbolized either as

$$\alpha_1^{p-i}\alpha_2^{i}\ \beta_1^{p-i}\beta_2^{j},$$

or as

$$\beta_1^{p-i}\beta_2^{i}\ \alpha_1^{p-j}\alpha_2^{j},$$

which mean exactly the same actual product $a_i a_j$. Conversely any product of p α's and p β's stands for a unique product of two coefficients a_i, a_j. The value of this result will appear in the sequel, for at present it seems artificial and useless.

In general, to express a product

$$a_0^q a_1^r \ldots a_p^t$$

of degree i in the coefficients of the ground form we introduce i equivalent symbols $\alpha, \beta, \gamma \ldots$: or, what is the same thing, we render it linear in i different coefficient sets $[a]$, $[b]$, $[c] \ldots$

by use of Aronhold operators $\left(b\dfrac{\partial}{\partial a}\right)$, $\left(c\dfrac{\partial}{\partial a}\right)\ldots$, and then substitute the symbols α, β, $\gamma\ldots$ as before.[1]

For example, if $f = (a_0, a_1, a_2, a_3 \between x_1, x_2)^3$ is a binary cubic, then

$$\begin{aligned} a_0 a_1{}^2 &= \alpha_1{}^3\,\beta_1{}^2\,\beta_2\,\gamma_1{}^2\,\gamma_2 \\ &= \beta_1{}^3\,\gamma_1{}^2\,\gamma_2\,\alpha_1{}^2\,\alpha_2 \\ &= \gamma_1{}^3\,\alpha_1{}^2\,\alpha_2\,\beta_1{}^2\,\beta_2. \end{aligned}$$

Conversely, a polynomial of degree three in each set α, β, γ symbolizes a polynomial of degree three in a_0, a_1, a_2, a_3. Thus

$$\lambda\alpha_1{}^2\alpha_2\beta_1\beta_2{}^2\gamma_2{}^3 + \mu\alpha_1\alpha_2{}^2\beta_1{}^2\beta_2\gamma_1\gamma_2{}^2 = \lambda a_1 a_2 a_3 + \mu a_1 a_2{}^2.$$

EXAMPLES

1. Prove that, if $(\alpha\beta)$ denote $(\alpha_1\beta_2 - \alpha_2\beta_1)$, the Jacobian of two quadratics $\alpha_x{}^2$, $\beta_x{}^2$ is symbolized by $4(\alpha\beta)\alpha_x\beta_x$.

2. If $\begin{vmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \\ c_0 & c_1 & c_2 \end{vmatrix}$ is the determinant whose rows are the coefficients of three quadratics symbolized by $\alpha_x{}^2$, $\beta_x{}^2$, $\gamma_x{}^2$, prove that this determinant is symbolized by $-(\beta\gamma)(\gamma\alpha)(\alpha\beta)$.

3. If α and β are equivalent symbols, prove that both $(\alpha\beta)\alpha_x\beta_x$ and $(\beta\gamma)(\gamma\alpha)(\alpha\beta)$ vanish identically.

4. The symbol $(\alpha\beta)\alpha_x{}^{p-1}\beta_x{}^{q-1}$ denotes the Jacobian of a binary p-ic, $\alpha_x{}^p$, and a binary q-ic $\beta_x{}^q$, neglecting a numerical factor pq.

5. Show that for a cubic $(a_0, a_1, a_2, a_3 \between x_1, x_2)^3$ symbolized by $\alpha_x{}^3$ and $\beta_x{}^3$, the Hessian is given by $(\alpha\beta)^2\alpha_x\beta_x$.

[The Hessian is $\begin{vmatrix} a_0x_1 + a_1x_2 & a_1x_1 + a_2x_2 \\ a_1x_1 + a_2x_2 & a_2x_1 + a_3x_2 \end{vmatrix}$. Use symbol α for row$_1$, and β for row$_2$].

6. Show that any determinant of order n whose rows (or columns) each involve the coefficients a_i of a quantic linearly can be symbolized by the use of n equivalent symbols, one for each row (or column).

---

[1] Various extensions of the use of these symbols, sometimes called umbral symbols, or umbræ, can be made, as, for example, in the next chapter. We pass over the question of *general forms*, whose order p may not be an integer; cf. *Encyklopädie der Math. Wiss.*, III, **3**, 6 (1922), p. 7: and also that of the relation of symbols to a power series

$$a_0 + a_1t + a_2t^2 + a_3t^3 + \ldots$$

whose coefficients are specific integers or combinations thereof. The reader will find a very interesting account of these by E. T. Bell, *Algebraic Arithmetic* (New York, 1927), 146–159.

CHAPTER XI

THE FIRST FUNDAMENTAL THEOREM

1. Symbolic Factors. Inner and Outer Products.

We are now in a position to consider a theorem of fundamental importance in the invariant theory. It enables us to construct, with the help of the symbols, as many invariants and covariants of given ground forms as we like; and conversely the theorem proves that by adhering to a specified mode of construction, *all* invariants and covariants may be found.

As a preliminary let us consider the Jacobian of three ternary quadratic forms U, V, W given by

$$\Sigma\, a_{ij} x_i x_j, \quad \Sigma\, b_{ij} x_i x_j, \quad \Sigma\, c_{ij} x_i x_j$$

$i, j = 1, 2, 3$, whose symbolic forms are, say,

$$\alpha_x^2, \quad \beta_x^2, \quad \gamma_x^2,$$

where $\alpha_x = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3$, which may be regarded as an inner product of the set α_1, α_2, α_3 and the set x_1, x_2, x_3. The symbolic form α_x^2 is in fact the square of a symbolic linear form or inner product. All such factors α_x, β_x, γ_x are called symbolic factors of the *first type* or symbolic inner products.

Now the Jacobian of U, V, W is

$$\begin{matrix} \dfrac{\partial U}{\partial x_1} & \dfrac{\partial V}{\partial x_1} & \dfrac{\partial W}{\partial x_1} \\[2ex] \dfrac{\partial U}{\partial x_2} & \dfrac{\partial V}{\partial x_2} & \dfrac{\partial W}{\partial x_2} \\[2ex] \dfrac{\partial U}{\partial x_3} & \dfrac{\partial V}{\partial x_3} & \dfrac{\partial W}{\partial x_3} \end{matrix} \qquad \begin{matrix} 2\alpha_x\alpha_1 & 2\beta_x\beta_1 & 2\gamma_x\gamma_1 \\ 2\alpha_x\alpha_2 & 2\beta_x\beta_2 & 2\gamma_x\gamma_2 \\ 2\alpha_x\alpha_3 & 2\beta_x\beta_3 & 2\gamma_x\gamma_3 \end{matrix}$$

since this determinant on expansion involves exactly two of each symbol α, β, γ in each term,

$$= 8 \begin{vmatrix} \alpha_1 & \beta_1 & \gamma_1 \\ \alpha_2 & \beta_2 & \gamma_2 \\ \alpha_3 & \beta_3 & \gamma_3 \end{vmatrix} \alpha_x \beta_x \gamma_x = 8(\alpha\beta\gamma)\alpha_x\beta_x\gamma_x,$$

which is the symbolic form of the Jacobian of three ternary quadratics. It involves a determinantal factor $(\alpha\beta\gamma)$ which is an *outer product* of the symbolic linear sets α, β, γ and sometimes is called a *bracket factor* or factor of the *second type*, to distinguish it from the numerical factor 8 and the symbolic *inner product* or linear factors α_x, β_x, γ_x which are of the *first type.*

Factors of these two kinds are characteristic of covariants (and invariants) expressed symbolically; indeed the fundamental theorem will demonstrate that every rational integral covariant of one or more ground forms involving variables $x_1, x_2, \ldots, x_n$ can be expressed symbolically entirely by means of these two kinds of symbolic factor.

Since an invariant contains no x, it is symbolized entirely by means of α, β, γ, and according to this theorem it cannot involve factors of type α_x. So it is composed entirely of the second type, the determinantal factor involving n symbols.

2. Effect of Linear Transformation on the Symbols.

First we must consider what happens to a general p-ic in n variables, symbolized by

$$(a_1x_1 + a_2x_2 + \ldots + a_nx_n)^p \equiv a_x^p \quad . \quad . \quad (1)$$

when a linear transformation $x \rightarrow x'$ is made. As in §6, p. 177, let the general term of this p-ic have coefficient

$$a_ia_ja_k \ldots = a_{ijk} \ldots . \quad . \quad . \quad . \quad . \quad (2)$$

If the linear transformation is

$$x_i = \xi_i x_1' + \eta_i x_2' + \ldots + \omega_i x_n', \qquad i = 1, 2, \ldots, n, \quad (3)$$

the effect on a_x is

$$(a_1\xi_1 + \ldots + a_n\xi_n)x_1' + (a_1\eta_1 + \ldots + a_n\eta_n)x_2' + \ldots \quad (4)$$

or
$$a_\xi x_1' + a_\eta x_2' + \ldots + a_\omega x_n'.$$

Hence, as in (11), p. 149, *the symbols* a *behave as a set contragredient to* x, a result of fundamental importance. Also the p-ic itself becomes

$$(a_\xi x_1' + \ldots + a_\omega x_n')^p \equiv a'_{x'}{}^p \quad . \quad . \quad . \quad (5)$$

say, where the general term of the p-ic in $x_1', \ldots, x_n'$ has coefficient

$$a'_{ijk}\ldots = a_i' a_j' a_k' \ldots = a_{\xi'} a_{\eta'} a_{\zeta'} \ldots, \quad . \quad . \quad (6)$$

$\xi', \eta', \zeta' \ldots$ denoting the ith, jth, kth $\ldots$ of the set $\xi, \eta, \zeta \ldots$.

For example, the cubic a_x^3 becomes $a'_{x'}{}^3$, and the coefficient of $x_1'^2 x_2'$ is $a'_{112} = a_1' a_1' a_2' = a_\xi^2 a_\eta$.

It will be seen that the new coefficients are obtained from the old by exhaustively polarizing the p-ic a_x^p with regard to ξ, η, $\zeta \ldots$ in all possible ways, in agreement with the previous result for binary forms. Similarly for other symbols $b, c, \ldots,$ belonging to general forms $b_x^q, c_x^r, \ldots$.

Next, the effect of the transformation on a symbolic determinant $(ab \ldots m)$ is to give

$$(a'b' \ldots m') = \begin{vmatrix} a_\xi & a_\eta & & a_\omega \\ b_\xi & b_\eta & & b_\omega \\ & & & \\ m_\xi & m_\eta & & m_\omega \end{vmatrix} = (\xi\eta \ldots \omega)(ab \ldots m), \quad (7)$$

another result of fundamental importance. Further, the effect on minors of this symbolic determinant is given by the theorem of corresponding matrices. Thus

$$\begin{aligned} a_1' = a_\xi,\ a_2' &= a_\eta, \ldots, a_n' = a_\omega \\ (a'b')_{ij} &= (ab \mid \xi'\eta') \\ (a'b'c')_{ijk} &= (abc \mid \xi'\eta'\zeta'), \text{ \&c.} \quad . \quad . \quad . \quad . \quad (8) \end{aligned}$$

3. Converse Theorem.

This at once gives us the converse of the First Fundamental Theorem in the following form:

Every symbolic product whose factors are solely of the two types $\mathrm{a_x}$ *or* (ab ... m) *satisfies the invariant condition.*

For let $\quad P = (abc \ldots)(def \ldots) \ldots g_x h_x \ldots$

be such a product, consisting of w bracket factors and ϖ x-factors, involving symbols $a, b, \ldots, g, h, \ldots$ of one or more given ground forms (1).

Let $P' = (a'b'c' \ldots)(d'e'f' \ldots) \ldots g'_{x'} h'_{x'} \ldots$ be the corresponding function of the accented symbols and variables.

Since $g'_{x'} = g_x$ and $(a'b'c' \ldots) = (\xi\eta\zeta \ldots)(abc \ldots)$, &c., we have

$$P' = (\xi\eta\zeta \ldots \omega)^w (abc \ldots)(def \ldots) \ldots g_x h_x \ldots \quad = |M|^w P$$

which proves the theorem.

As an example, consider the ternary symbolic product

$$P = (abc)(abd)\, c_x d_x,$$

so that

$$P' = (a'b'c')(a'b'd')\, c'_{x'} d'_{x'} = (\xi\eta\zeta)^2 (abc)(abd)\, c_x d_x.$$

Let this be expanded in full, as a polynomial in all the arguments $a_1', \ldots, d_3', x_1', x_2', x_3', a_1, \ldots, x_3$. On the left, a typical term is

$$a'^2_1\, b'^2_2\, c'_3\, d'_3\, c'_1\, d'_1\, x'^2_1,$$

which is of degree two in each set a', b', c', d', and x'. Since each symbolic factor of either kind $(a'b'c')$ or $c'_{x'}$ is linear and homogeneous in its sets a', b', c', x', it is at once apparent that the dimensions of the typical term are also given merely by counting how many of each symbol a' or variable x' occur in the *unexpanded* product P': and this will be true in general.

Thus P is a quadratic in x because it has two factors c_x, d_x, and further, if P is not a mere symbolic covariant but an actual covariant, we infer that the symbols a, b, c, d refer to *quadratic* ground forms, because there are *two* of each symbol.

Hence P is a covariant of weight two, the index of $(\xi\eta\zeta)$, and of degree four, in one quadratic, when

$$a_x^2 = b_x^2 = c_x^2 = d_x^2$$

are equivalent symbolic forms: or again is of partial degrees (1, 3) in two quadratics a_x^2 and $b_x^2 = c_x^2 = d_x^2$ where symbols b, c, d but not a refer to the second quadratic: or again is of partial degrees (2, 1, 1) in three quadratics: and so on.

4. The Valency Condition $pq = nw + \varpi$ for Single Ground Form.

If we continue to illustrate with the ternary case of the p-ic

$$f = a_x^p = b_x^p = c_x^p = \dots$$

where a, b, c are equivalent symbols, we can see at once that a product such as

$$P = (abc)\,(bcd)\,(cda)\,(dab)\dots$$

involving w bracket factors but no x-factors contains $3w$ symbols in all, each distinct symbol, such as a, occurring p times. Hence for a ternary invariant involving q different symbols of a p-ic, the relation

$$pq = 3w$$

is essential. In general the condition

$$pq = nw$$

connects the degree q and the weight w of any invariant of a p-ic in n variables. For the subsequent proof of the fundamental theorem will show that any invariant can be expressed as a sum of symbolic products such as P.

Again, a more general product, say

$$Q = (abc)\,(bcd)\,(cda)\,a_x^2\,b_x^2\,c_x\,d_x^2$$

involving ϖ x-factors and w bracket factors, is a covariant of an n-ary p-ic if

$$pq = nw + \varpi,$$

as is seen by counting all the symbols a, b, c ... which occur. Here $\varpi = 7$, $w = 3$, $n = 3$, $p = 4$, $q = 4$.

It is desirable to give a name to such relations between positive integers indicating degree, weight, or order of invariants and the like. Let them be called *valency conditions.*

EXAMPLES

1. A binary form of even order has no covariant of odd order.

2. A binary quartic has at least two invariants, $(\alpha\beta)^4$, $(\beta\gamma)^2(\gamma\alpha)^2(\alpha\beta)^2$. Express these non-symbolically when

$$(a_0, a_1, a_2, a_3, a_4 \between x_1, x_2)^4 = \alpha_x^4 = \beta_x^4 = \gamma_x^4.$$

3. A binary *n-ic* has an invariant of degree two only if n is even. For let $(a_0, a_1, \dots, a^n \between x_1, x_2)^n = \alpha_x{}^n = \beta_x{}^n$. Then

$$\begin{aligned}(\alpha\beta)^n &= a_0 b_n - n a_1 b_{n-1} + \binom{n}{2} a_2 b_{n-2} - \dots + (-)^n a_n b_0 \\ &= 0 \text{ if } n \text{ is odd.}\end{aligned}$$

5. First Fundamental Theorem for a System of Linear Forms.

Let us first prove the theorem for linear forms in n homogeneous variables. Starting with a set of variables $x_1, x_2, \dots, x_n$ and any number $q (\geqslant n)$ of linear forms, not necessarily all distinct,

$$\left.\begin{aligned} A &= a_x = a_1 x_1 + a_2 x_2 + \dots + a_n x_n \\ B &= b_x = b_1 x_1 + b_2 x_2 + \dots + b_n x_n \\ & \\ K &= k_x = k_1 x_1 + k_2 x_2 + \dots + k_n x_n \end{aligned}\right\rceil \qquad (9)$$

we enunciate the theorem as follows.

Fundamental Theorem for Linear Forms.

Every rational integral projective invariant of linear forms $a_x, b_x, c_x, \dots, k_x$ *is expressible as an aggregate of terms consisting entirely of bracket factors of the type* $(ab \dots m)$ *together with numerical coefficients.*

This type of factor $(ab \dots m)$ is indeed an n-rowed determinant of the coefficient matrix

$$\Phi = \begin{array}{cccc} a_1 & a_2 & & \\ b_1 & b_2 & & \\ & & & \\ \lfloor k_1 & k_2 & \dots & k_n \rfloor \end{array} \qquad . \quad . \quad . \quad (10)$$

Proof.—

Let the result of a linear transformation $x \longrightarrow x'$ on the variables change the linear forms A to $a'_{x'}$, B to $b'_{x'}$, &c. Then the new coefficient matrix is

$$\Psi = \begin{array}{cccccccc} a_1' & a_2' & \dots & a_n & \quad a_\xi & a_\eta & & \\ . & . & . & . & . & . & & \\ \lfloor k_1' & k_2' & \dots & k_n' \rfloor & \quad \lfloor k_\xi & k_\eta & \dots & k_\omega \rfloor \end{array}$$

where $$a_\xi = a_1 \xi_1 + a_2 \xi_2 + \dots + a_n \xi_n, \qquad (11)$$

and, as before, the coefficient matrix of the linear transformation is

$$M = \begin{bmatrix} \xi_1 & \eta_1 & \zeta_1 & \dots & \omega_1 \\ \xi_n & \eta_n & \zeta_n & \dots & \omega_n \end{bmatrix} \qquad (12)$$

Now if $I(a, b, \dots, k)$ is an invariant rational and integral in the nq coefficients of these forms, then by hypothesis

$$I(a', b', \dots, k') = |M|^w I(a, b, \dots, k). \quad . \quad (13)$$

This is an identity in the n^2 elements $\xi_1, \dots, \omega_n$, so that it still remains an identity after differentiation by $\frac{\partial}{\partial \xi_i}$ or more generally by the Cayley operator (§2, p. 113)

$$\Omega = \begin{matrix} \frac{\partial}{\partial \xi_1} & \frac{\partial}{\partial \xi_2} & & \frac{\partial}{\partial \xi_n} \\ \\ \frac{\partial}{\partial \omega_1} & \frac{\partial}{\partial \omega_2} & & \frac{\partial}{\partial \omega_n} \end{matrix} \qquad (14)$$

Regarded as a function of $\xi_1, \dots, \omega_n$ the right member of (13) is a polynomial, homogeneous and of degree w in the set $\xi_1, \xi_2, \dots, \xi_n$ as well as in each of the other sets η or $\zeta \dots$ or ω. For such sets only enter by way of the determinant $|M|$ which is linear in each set. To balance this the left member of (13) must also be homogeneous and of degree w in each set. But since $I(a', b', \dots, k')$ is explicitly a polynomial in $a_1', \dots, k_n'$, which are the same as $a_\xi, \dots, k_\omega$, it follows that every term on the left of (13) contains exactly w factors $a_\xi, b_\xi \dots$ involving ξ, w factors $a_\eta, b_\eta \dots$ involving η, and so on.

We now operate on both sides of (13) with Ω. Since ((33), p. 122),

$$\Omega a_\xi b_\eta c_\zeta \dots m_\omega r_\xi \dots = \Sigma (abc \dots m)\, r_\xi \dots, \qquad (15)$$

where in Σ the letters $a, b, \dots, m, r, \dots$ are suitably permuted, the precise way being immaterial, we obtain for every term on the left an aggregate of terms each containing a bracket factor like $(abc \dots m)$ but one ξ fewer, one η fewer, &c. On the right we obtain (§3, p. 114)

$$w(w+1) \dots (w+n-1)\, |M|^{w-1} I(a, b, \dots, k). \qquad (16)$$

Repeating this process until Ω operates w times on the original identity, we exhaust all of $\xi, \eta, \ldots, \omega$ on either side, gaining at every stage one new bracket factor in each term on the left, and thereby accounting for all the letters $a, b, c, \ldots$ on the left. Thus

$$\Sigma\lambda(abc\ldots m)(rs\ldots)\ldots = \mu I(a, b, c, \ldots), \quad . \quad (17)$$

where λ, μ are numerical constants, and $\mu \neq 0$ (Ex. **5**, p. 115). Dividing by μ we finally express $I(a, b, c, \ldots)$ in the desired form.

Corollary I.—*The number* q *of linear forms, including repetitions, in an invariant is a multiple of* n. In fact $q = nw$.

Corollary II.—*The simplest invariant of linear forms* a_x, b_x ... *is* (ab ... m) *involving* n *different forms*; *it is the determinant of the coefficient matrix of* n *such forms.*

For it is the simplest expression of the requisite type, and it would vanish if two of the sets a, b in the determinant were identical.

Corollary III.—*No invariant exists of less than* n *linear forms.*

Corollary IV.—*Each term in the summation on the left of* (17) *is an invariant.* This follows from §3, p. 184.

The above proof contains the leading idea required in the general proof for invariants of any ground forms. It has the great merit of carrying with it the actual method for throwing a given invariant into this convenient symbolic form. And although the process would be tedious in complicated cases, the reader will find it very instructive to follow out the steps in detail for a few simple cases, in order to grasp the several principles of the proof.

6. Invariants of One or More General Ground Forms.

Next we prove the Fundamental Theorem for the case of invariants of forms of higher order than the linear. But for the initial step, the proof is exactly the same. Thus let there now be q given ground forms of orders $r, s, \ldots, t$ respectively, written symbolically as

$$\left.\begin{array}{l} A = a_x{}^r = (a_1x_1 + a_2x_2 + \ldots + a_nx_n)^r \\ B = b_x{}^s = (b_1x_1 + b_2x_2 + \ldots + b_nx_n)^s \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ K = k_x{}^t = (k_1x_1 + k_2x_2 + \ldots + k_nx_n)^t \end{array}\right\}, \quad . \quad (18)$$

where as before these q forms need not necessarily all be distinct. We utilize, in all, qn different symbols $a_1, a_2, \ldots, k_n$: a product of r a's gives an actual coefficient of the first ground form A; a product of s b's, one belonging to the next form B; and so on.

In this way we express the first ground form as a perfect rth power of a symbolic linear form; and similarly for the rest.

Then if the linear transformation be made as before, $a_1x_1 + \ldots + a_nx_n$ becomes $a_\xi x_1' + a_\eta x_2' +$ &c. Accordingly the form A of order r becomes

$$(a_\xi x_1' + a_\eta x_2' + \ldots)^r = (a_1'x_1' + a_2'x_2' + \ldots)^r \equiv a'_{x'}{}^r,$$

where a' is the corresponding symbol associated with the new variable x'. Similarly for the symbols $b, \ldots, k$.

Now if the invariant $I(a, b, \ldots, m)$ is a polynomial in the actual coefficients of forms $A, B, \ldots, K$, it is also a polynomial in the qn symbols $a_1, \ldots, k_n$, when actual coefficients are resolved into symbols as in §5 (17). It can then be treated exactly as in the previous case; and the result of the operation Ω^w upon the identity

$$I(a', \ldots) = |M|^w I(a, \ldots)$$

expresses the invariant entirely in terms of the bracket factors $(ab \ldots m)$.

Since this bracket factor on expansion is linear in the set $a_1, \ldots, a_n$, and therefore vanishes if a appears twice in the same factor, we infer that exactly r such factors each contain a symbol a; s factors contain b; and so on.

Further there are exactly w such factors, one for each step in the operation Ω^w (cf. Ex. **1**, **4**, p. 123). Hence the total number of symbols $a, \ldots, k$ in one such product of factors is nw, so that the following valency relation must hold:

$$r + s + t \ldots = nw.$$

Corollary I.—*The degree* q *and the weight* w *of an invariant of a single ground form of order* p *in* n *variables satisfy the condition* pq = nw.

For in this case $r = s = t = \ldots = p$, and the q forms A, B, ... are all the same. Symbolically we denote such a form f as

$$f = a_x^p = b_x^p = \ldots = k_x^p.$$

7. Examples of Invariants. Interchange of Equivalent Symbols.

Various examples have already been given of the symbolic form of invariants. We must now consider more particularly a few cases where a single ground form is involved and q equivalent symbols are needed for expressing an invariant of degree q in its coefficients.

Let the ground form be

$$f = a_x^p = b_x^p = c_x^p = \text{, \&c.}$$

Example 1.—

If f is a binary quadratic,

$$f = a_x^2 = a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2,$$

then its discriminant is $D = a_{11}a_{22} - a_{12}^2$, which is symbolized as $a_1^2b_2^2 - a_1a_2b_1b_2$, or equally well as $b_1^2a_2^2 - b_1b_2a_1a_2$, which is formed from the first expression by interchanging the rôles of a and b throughout. This process is called the *interchange of equivalent symbols*, and for two symbols introduces two alternative forms of the expression, for three symbols, 3! alternatives. and for n symbols, n! alternatives. when interchanges are made in all possible ways. Taking the first symbolic form, we have

$$D = a_1b_2(a_1b_2 - a_2b_1) = a_1b_2(ab).$$

Interchanging equivalent symbols, we also have

$$D = b_1a_2(ba).$$

But $(ba) = -(ab)$. Hence by addition

$$2D = a_1b_2(ab) - b_1a_2(ab) = (ab)^2.$$

So $$D = \tfrac{1}{2}(ab)^2.$$

This little device which is of great use in throwing invariants into convenient form deserves close study. Let us now adopt a previous notation, in order to abbreviate such a process. So in this example we write

$$D = a_1 b_2 (ab), \qquad 2D = \dot{a}_1 \dot{b}_2 (ab) = (ab)(ab) = (ab)^2.$$

Example 2.—

The discriminant of a ternary quadratic $f = a_x^2 = \Sigma\, a_{ij} x_i x_j$ is an invariant of degree three in the coefficients a_{ij}:

$$D = |a_{ij}| = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} \qquad (a_{ij} = a_{ji}).$$

Here $a_{ij} = a_i a_j = b_i b_j = c_i c_j$, and three equivalent symbols are needed to represent the expansion of the three-rowed determinant D. Since D is linear in its column elements we may write (Ex. 6, p. 181)

$$D = \begin{vmatrix} a_1^2 & b_1 b_2 & c_1 c_3 \\ a_2 a_1 & b_2^2 & c_2 c_3 \\ a_3 a_1 & b_3 b_2 & c_3^2 \end{vmatrix} = a_1 b_2 c_3 \begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{vmatrix}$$

after extracting common factors from columns. Accordingly

$$D = a_1 b_2 c_3 (abc).$$

Now we interchange the equivalent symbols a, b, c in all six possible ways, and add the results. Since $(abc) = -(acb) = (bca)$, &c., this gives

$$\begin{aligned} 6D &= (a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1)(abc) \\ &= \dot{a}_1 \dot{b}_2 \dot{c}_3 (abc) = (abc)^2. \end{aligned}$$

So $$D = \tfrac{1}{6}(abc)^2;$$

and by the fundamental theorem this is an invariant.

In general the discriminant of a quadratic form in n variables requires n equivalent symbols $a, b, \ldots, m$, and is

$$D = |a_{ij}| = \frac{1}{n!}(abc \ldots m)^2.$$

Example 3.—

The ternary cubic $f = a_x^3 = \Sigma\, a_{ijk} x_i x_j x_k$ has two invariants which can be symbolized by

$$S = (abc)\,(abd)\,(acd)\,(bcd)$$
$$T = (abc)\,(abd)\,(ace)\,(bcf)\,(def)^2.$$

Here a, b, c, d, e, f are equivalent symbols whose actual coefficient sets are all equal:

$$a_{ijk} = a_i a_j a_k = b_i b_j b_k = \ldots = f_i f_j f_k, \qquad i, j, k = 1, 2, 3.$$

Merely counting the number of a's, &c., verifies the invariant property of S and T; counting bracket factors gives the weight w, and the number of different letters a, b, &c., gives the degree q. Thus

$$S : n = 3, \quad p = 3, \quad q = 4, \quad w = 4,$$
$$T : n = 3, \quad p = 3, \quad q = 6, \quad w = 6,$$

which naturally satisfy the valency condition $pq = nw$.

Symbolic invariants *may* vanish identically: thus the simpler looking invariant $(abc)^3$ is zero. For

$$(abc)^3 = (bac)^3$$

on interchanging equivalent symbols. But

$$(bac)^3 = [-(abc)]^3 = -(abc)^3; \text{ hence } (abc)^3 + (abc)^3 = 0.$$

Example 4.—

$(abc \ldots m)^p$ is the invariant of lowest weight and degree for the general form of order p, but it vanishes identically if p is odd.

8.[1] Double Convolution of Symbols referring to a Quadric Form.

The following theorem follows up the process used in the above Example 2. It gives in general what was originally discovered by Gordan in his successful researches upon quadratics in ternary and quaternary forms. The technique has already been explained in §9, p. 44.

---

[1] This section 8 may be omitted on a first reading.

THEOREM.—*A symbolic product* P *of bracket factors, wherein* h *equivalent symbols of a quadric are explicitly convolved once, is expressible as a sum of terms in which these symbols are explicitly convolved twice.*

Proof.—

For clearness we first consider a simple example, $n = 5$, $h = 3$. Let

$$P = (abc\,rs)\,(ab\,M)\,(cN)\,Q,$$

where a, b, c are equivalent symbols, convolved once. Since they belong to a quadric, they have duplicates, which occur elsewhere in factors of P, either convolved in one factor or not. In this example they occur in two factors, and the remaining symbols r, s, M, N, Q represent suitable arbitrary elements, involving neither a, nor b, nor c. This last condition is essential to the success of the proof.

Interchanging a, b, c in all 3! ways we have 3! alternative expressions for P. We add them together and obtain

$$3!\,P = (abc\,rs)\,(abM)\,(cN)\,Q + (bac\,rs)\,(baM)\,(cN)\,Q + \ldots.$$

But the first factor of each term is of type $\pm\,(abc\,rs)$. Hence

$$\begin{aligned} 3!\,P &= (abc\,rs)\,\{(abM)\,(cN)\,Q - (baM)\,(cN)\,Q + \ldots\} \\ &= 2(abc\,rs)\,\{(abM)\,(cN)\,Q + (bcM)\,(aN)\,Q + (caM)\,(bN)\,Q\} \\ &= 2(abc\,rs)\,(\dot{a}\dot{b}M)\,(\dot{c}N)\,Q \\ &= 2(abc\,rs)\,(abc\,\dot{m}'\dot{m}'')\,(\dot{m}N)\,Q, \end{aligned}$$

where $M = mm'm''$, by a fundamental identity I, §9, p. 45. This proves the theorem for P. Here every step has been given because each is typical of the general case. Thus we take

$$P = (A_iB_jK)\,(A_iM)\,(B_jN)\,Q,$$

where A_i denotes i symbols $a_1, a_2, \ldots, a_i$, and B_j, $b_1, b_2, \ldots, b_j$, all $i + j$ of which are equivalent and refer to one quadric. Then by all possible interchanges

$$\begin{aligned} (i+j)!\,P &= (A_iB_jK)\,\{\Sigma \pm (A_iM)\,(B_jN)\,Q\} \\ &= i!\,j!\,(A_iB_jK)\,(\dot{A}_iM)\,(\dot{B}_jN)\,Q. \end{aligned}$$

If now the $n - i$ columns M are written $M_j M'_{n-i-j}$, we have by a fundamental identity

$$(i+j)!\ P = i!\ j!\ (A_i B_j K)\,(A_i B_j \dot{M}'_{n-i-j})\,(\dot{M}_j N)\,Q,$$

with $A_i B_j$ convolved twice. Similarly if the first factor of P has $A_i B_j C_k \ldots$ all referring to the one quadric, we obtain by the same process

$$(i+j+k+\ldots)!\ P = i!\ j!\ k! \ldots (A_i B_j C_k \ldots)\,(A_i B_j C_k \ldots) \ldots Q.$$

This proves the theorem.

Corollary.—*The quadric may be replaced by a form of order* 2p *in which* P *contains equivalent symbols convolved* 2p — 1 *times. The process produces a further convolution.*

Again, the new convolution may be assembled in any assigned bracket factor g except the first. The process deranges the other original symbols of the factor g, but leaves them *implicitly* convolved.

For instance, in the first example

$$g = (abM) = (abmm'm''),$$

and after convolving abc in g, the symbols M, expressed in the currency of a as m, m', m'', are implicitly convolved in a series of three terms $\dot{m}\dot{m}'$, $\dot{m}''$. But the factor g selected might equally well have been (cN) or even a factor of Q, since the fundamental identities apply in each case.

9. Solution of Symbolic Linear Equations.

Equations frequently arise which are linear in one set x of variables. Even when expressed symbolically they can be solved by the ordinary methods. Thus for a quaternary case consider

$$a_x a_y = 0, \quad b_x b_z = 0, \quad c_x c_t = 0, \qquad . \quad . \quad (19)$$

which arise from three quadrics polarized with regard to y, z, t respectively (§7, p. 177).

Writing each in full as far as x is concerned we have $a_x a_y = a_y(a_1 x_1 + a_2 x_2 + a_3 x_3 + a_4 x_4)$, and so on. Then after multiplying the three by $b_z c_t (bc)_{23}$, $c_t a_y (ca)_{23}$, $a_y b_z (ab)_{23}$ respectively and adding we obtain

$$a_y b_z c_t \big((abc)_{123} x_1 + (abc)_{423} x_4 \big) = 0. \quad . \quad (20)$$

Solving this and similar results we find

$$\begin{aligned} x_1 : x_2 : x_3 : x_4 = a_y b_z c_t (abc)_{234} : a_y b_z c_t (abc)_{314} : \\ : a_y b_z c_t (abc)_{124} : a_y b_z c_t (abc)_{213}. \quad . \quad . \quad (21) \end{aligned}$$

Now if we introduce arbitrary values u_1, u_2, u_3, u_4 such that $u_x = \Sigma\, u_i x_i = 0$, then

$$(abcu)\, a_y b_z c_t = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad (22)$$

It will be seen that the above equations can be solved exactly as linear equations $a_x = 0$, $b_x = 0$, $c_x = 0$, provided that we maintain, throughout, the original common factors a_y, b_z, c_t which give actuality to the equations.

Further, if a, b, c are *equivalent* symbols we can follow the methods of p. 192. Thus (21) and (22) become

$$x_1 : x_2 : x_3 : x_4 = (abc \,|\, yzt)\, (abc)_{234} : \text{ \&c.} \quad . \quad (23)$$

and

$$(abcu)\, a_y b_z c_t = \tfrac{1}{6} (abcu)\, (abc \,|\, yzt). \quad . \quad . \quad (24)$$

CHAPTER XII

MULTILINEAR FORMS

1. Multilinear Forms.

We now set about proving the Fundamental Theorem in a more general form, so as to include covariants (§12, p. 144) as well as invariants within its scope. To do this we must examine more in detail other possible types of ground form than that which depends only on one set of variables $x_1, x_2, \ldots, x_n$. One remarkable feature of the symbolic method is that the more general cases are as easy to handle as the special cases, as we saw in considering the invariant, multilinear in q ground forms of order p, which is symbolically of the same structure as an invariant of degree q in one ground form.

Let us consider the ground form $a_x{}^p$, of order p in n homogeneous variables x, from this point of view. We take

$$f = a_x{}^p = \Sigma\, a_{ijk\ldots}\, x_i x_j x_k \ldots, \tag{1}$$

where each of $i, j, k \ldots$ has values $1, 2, \ldots, n$. If we polarize this with regard to $p-1$ sets $y, z, \ldots$ in succession we render it multilinear in p sets $x, y, z \ldots.$ For brevity of statement we consider the ternary cubic ($n = p = 3$). Thus

$$a_x{}^3 = (a_1x_1 + a_2x_2 + a_3x_3)^3 = \Sigma\, a_{ijk}\, x_i x_j x_k, \tag{2}$$

where $\quad a_i a_j a_k = a_{ijk} = a_{jik} = \&\text{c.} \quad \ldots\ldots\ldots \quad (3)$

Operating with

$$\left(y_1 \frac{\partial}{\partial x_1} + y_2 \frac{\partial}{\partial x_2} + y_3 \frac{\partial}{\partial x_3}\right) \equiv \left(y \frac{\partial}{\partial x}\right)$$

followed by $\quad \left(z_1 \dfrac{\partial}{\partial x_1} + z_2 \dfrac{\partial}{\partial x_2} + z_3 \dfrac{\partial}{\partial x_3}\right) \equiv \left(z \dfrac{\partial}{\partial x}\right),$

where sets y and z are independent of x, we obtain

$$\frac{1}{3!}\left(z \frac{\partial}{\partial x}\right)\left(y \frac{\partial}{\partial x}\right) a_x{}^3 = \tfrac{1}{2}\left(z \frac{\partial}{\partial x}\right) a_x{}^2 a_y = a_x a_y a_z = \Sigma\, a_{ijk}\, x_i y_j{}^z$$

Such is called a trilinear ternary form: it is multilinear in three sets of variables. A simpler example of multilinearity is the bilinear form in two sets x, y, say

$$a_x a_y = \Sigma\, a_{ij}\, x_i y_j, \qquad (a_{ij} = a_{ji}),$$

arising as first polar of the quadratic a_x^2.

Owing to the symmetrical suffix relations (3), such multilinear forms are not the most general; howbeit they serve as an introduction to the general form.

Definition of General Multilinear Form in Sets of n Variables. —*The form* $\Sigma\, \mathrm{a}_{\mathrm{ijk}\ldots}\, \mathrm{x}_\mathrm{i} \mathrm{y}_\mathrm{j} \mathrm{z}_\mathrm{k} \ldots$ *is the general multilinear form in* p *sets of* n *variables, if every derangement of the suffixes* ijk . . . *of the typical coefficient alters its value.*

For instance, the general ternary bilinear form is

$$\Sigma\, a_{ij}\, x_i y_j = \begin{cases} \quad a_{11} x_1 y_1 + a_{12} x_1 y_2 + a_{13} x_1 y_3 \\ + a_{21} x_2 y_1 + a_{22} x_2 y_2 + a_{23} x_2 y_3 \\ + a_{31} x_3 y_1 + a_{32} x_3 y_2 + a_{33} x_3 y_3, \end{cases}$$

where $a_{ij} \neq a_{ji}$. Such a form cannot be derived by polarizing a quadratic $\Sigma\, a_{ij}\, x_i x_j$, since the coefficients of $x_1 y_2$ and $x_2 y_1$ differ.

2. Symbolic Representation of Multilinear Forms.

Following the lines suggested in the case of the binary form we readily find a suitable symbol for the general form in p sets $x, y, z \ldots,$

$$f = \Sigma\, a_{ijk\ldots}\, x_i y_j z_k \ldots.$$

Operating on f with $\dfrac{\partial^p}{\partial x_i \partial y_j \partial z_k \ldots}$ we obtain the single coefficient $a_{ijk\ldots}$. Then if we write

$$\frac{\partial^p f}{\partial x_i \partial y_j \partial z_k \ldots} = a_i b_j c_k \ldots, \qquad (i, j, k \ldots = 1, 2, 3, \ldots, n),$$

as in (17), p. 173, the p symbols a_i, b_j, c_k . . . commute with one another, because of the fundamental law of partial differentiation. Also they have an actual significance if they occur in a product involving exactly p of them—one a, one b, and so on.

Hence

$$a_{ijk\ldots} = a_i b_j c_k \ldots = b_j a_i c_k = \;\; \&\text{c.};$$

so that

$$\begin{aligned} f &= \Sigma a_i b_j c_k \ldots x_i y_j z_k \ldots \\ &= (a_1 x_1 + \ldots + a_n x_n)(b_1 y_1 + \ldots b_n y_n) \times \ldots \\ &= a_x b_y c_z \ldots , \end{aligned}$$

which symbolizes the general multilinear form. It involves p sets of variables (np variables in all) and p sets of symbols $a, b, c \ldots$ (np symbols in all). In particular if the form is symmetrical in two sets of variables x and y, we may take the corresponding symbols as equal, now writing

$$a_{ijk\ldots} = a_i a_j c_k \ldots , \qquad f = a_x a_y c_z \ldots ,$$

for in this case there is no reason to distinguish between $a_{ijk\ldots}$ and $a_{jik\ldots}$, that is to say, between $a_i b_j$ and $b_i a_j$.

By making some of the sets x, y, z equal we include as a special case of the multilinear form the form of higher order in fewer sets of n variables. For instance

$$a_x^2 b_z^2, \quad a_x a_y b_z b_t, \quad a_x c_y b_z d_t$$

are respectively (2, 2), (1, 1, 1, 1), (1, 1, 1, 1) forms. The (2, 2) form is a special case of the next form when $x = y$, $z = t$. This in turn is a special case of the third form, for it is symmetrical in x and y and in z and t. The last is the general quadrilinear form.

3. Classification of Multilinear Forms.

From the point of view of the invariant theory, the natural way to analyse these forms is according to the behaviour of the sets of variables x, y, ... To avoid constant repetition let us understand by the simple phrase: "the variables $x, y, \ldots$", the sets of variables $(x_1, x_2, \ldots, x_n)$, $(y_1, y_2, \ldots, y_n)$, &c. Then supposing all variables to undergo linear transformations, say $x \rightarrow x'$, $y \rightarrow y'$, ..., everything turns on whether these transformations are independent or not. For our present purpose we adopt the following classification in ascending order of complexity.

I. Variables all cogredient.
II. Variables cogredient and contragredient.
III. Variables independent.

If $x, y, z \dots$ are cogredient and $u, v, w \dots$ are all contragredient to $x, y, z \dots$, then I is really a special case of II, when one type of variable, say u, v, w, is entirely absent. The more general case III is at present dismissed, so that our chief concern is with II.

Remembering that x is a single column matrix, we write in matrix notation ((15), p. 149)

$$x = Mx', \quad y = My', \quad z = Mz', \quad \&c.,$$

for the cogredient transformations of coefficient matrix M; and

$$u' = uM, \quad v' = vM, \quad w' = wM, \quad \&c.,$$

for the induced transformations of the contragredient sets $u, v, w \dots$, where each set is a single row matrix.

4. Cogredient and Contragredient Symbols.

The following convention will now prove to be useful when we are concerned with multiple forms involving several sets of variables $x, y, \dots, u, v, \dots$. *We use the italics a, b, c as symbols associated with x, y, z, and Greek letters α, β, γ as symbols associated with the contragredient variables u, v, w.* Further, we write the symbolic inner product involving u and α as

$$\alpha_1 u_1 + \alpha_2 u_2 + \dots + \alpha_n u_n = u_\alpha.$$

Thus the general multilinear form in these variables is symbolized by

$$f = a_x b_y c_z \dots u_\alpha v_\beta w_\gamma \dots.$$

Upper and Lower Indices.

Non-symbolically the general coefficient of f involving p sets of variables $x, y, z \dots$ has already been written as $a_{ijk\dots}$ with p lower suffixes. But it has been found convenient to adopt upper indices, as in the theory of determinants, when the contragredient variables are involved, and to write

$$a_{ijk\dots}^{rst\dots}$$

for the typical coefficient of the multilinear form, involving p lower and q upper suffixes, answering to p sets of variables $x, y, z \dots$ and q sets of $u, v, w \dots$.

Definition of Tensor of Rank *r.—The set of coefficients* $\mathbf{a}_{\mathbf{ijk}\dots}^{\mathbf{rst}\dots}$ *with* p *lower and* q *upper suffixes, each taking all values*

1, 2, . . . , n *is called a tensor of orders* (p, q). It is sometimes called a tensor of rank $r(= p + q)$.

This use of the word *rank* must not be confused with its use in the theory of matrices. It has crept in through the work of physicists in the theory of relativity. For greater clearness let us call such a set *a tensor of orders* (p, q).

Examples.—

$a_1x_1 + a_2x_2 + a_3x_3 = \Sigma a_i x_i = a_x$ is the ternary linear form. It has orders (1, 0).

$u_1\alpha_1 + u_2\alpha_2 + u_3\alpha_3 = \Sigma u_i \alpha_i = u_\alpha$ is also a ternary linear form, but of orders (0, 1).

$\Sigma a_i^j x_i u_j = a_x u_\alpha$ is a bilinear form with contragredient variables x, u.

$\Sigma a_{ij} x_i y_j = a_x b_y$ is a bilinear form with cogredient variables, and so also is $\Sigma a^{ij} u_i v_j = u_\alpha v_\beta$.

$\Sigma a_{ij}^{kl} x_i x_j u_k u_l = a_x^2 u_\alpha^2$ is a (2, 2) form quadratic in two sets of contragredient variables x and u.

5. Equivalent Symbols.

Any homogeneous polynomial in the coefficients of such forms can be expressed symbolically by introducing equivalent symbols, as at p. 179. So a second degree polynomial in the coefficients of the (1, 1) form $a_x u_\alpha$ would require two sets of symbols, say

$$a_i^j = a_i \alpha_j = a_i' \alpha_j',$$

so that, for example,

$$a_1^{\,1} a_2^{\,3} = a_1 \alpha_1 a_2' \alpha_3' = a_1' \alpha_1' a_2 \alpha_3.$$

The general form $a_x b_y \ldots k_z u_\alpha v_\beta \ldots w$ involving p sets $x, y, \ldots$ and q sets $u, v \ldots$ has p symbol sets $a, b, \ldots, k$ and q symbol sets $\alpha, \beta, \ldots, \kappa$. Equivalent symbols $a', b', \ldots, k', \alpha', \ldots, \kappa'$ whenever used are such that the substitution

$$\begin{pmatrix} a' & b' & \ldots & k' & \alpha' & \beta' & \ldots & \kappa' \\ a & b & \ldots & k & \alpha & \beta & \ldots & \kappa \end{pmatrix}$$

leaves the actual form unchanged.

6. Effect of Linear Transformation on the Symbols.

We can now prove the following useful theorem.

Under linear transformation, M, M'^{-1}, *the symbols* a, b, c . . . *associated with variables* x *are cogredient with* u, *and symbols* α, β, γ . . . *are cogredient with* x.

For let $x \rightarrow x'$, $u \rightarrow u'$ denote the linear transformations. Then the general form in variables $x, y, \ldots, u, v \ldots$ retains its same orders when expressed in terms of $x', y', \ldots, u', v' \ldots$. If we use accents for symbols after transformation, we have in the case of the (1, 1) form

$$a_x u_\alpha = \Sigma\, a_i^j x_i u_j = \Sigma\, a'^j_i x_i' u_j' = a'_{x'} u'_{\alpha'}. \qquad . \quad (4)$$

This is directly and explicitly secured by taking

$$a'_{x'} = a_x, \quad u'_{\alpha'} = u_\alpha \qquad . \quad . \quad . \quad . \quad . \quad (5)$$

in all cases, and defining the new symbols a', α' by these relations. But they are the characteristic conditions of contragredience. Hence the transformation $a \rightarrow a'$ is contragredient to $x \rightarrow x'$, and $\alpha \rightarrow \alpha'$ to $u \rightarrow u'$. But u is contragredient to x. So a is cogredient with u, and α with x. Similarly for the general form.

In greater detail, let

$$\left.\begin{aligned} x_1 &= \xi_1 x_1' + \eta_1 x_2' + \zeta_1 x_3' \\ x_2 &= \xi_2 x_1' + \eta_2 x_2' + \zeta_2 x_3' \\ x_3 &= \xi_3 x_1' + \eta_3 x_2' + \zeta_3 x_3' \end{aligned}\right\}, \qquad . \quad . \quad . \quad (6)$$

where for shortness ternary variables are considered. Then the contragredient transformation for $u' \rightarrow u$ is (p. 148, (8))

$$\left.\begin{aligned} u_1' &= \xi_1 u_1 + \xi_2 u_2 + \xi_3 u_3 = u_\xi \\ u_2' &= \eta_1 u_1 + \eta_2 u_2 + \eta_3 u_3 = u_\eta \\ u_3' &= \zeta_1 u_1 + \zeta_2 u_2 + \zeta_3 u_3 = u_\zeta \end{aligned}\right\}. \qquad . \quad . \quad . \quad (7)$$

Consequently, for the symbols a, which are cogredient with u,

$$a_1' = a_\xi, \quad a_2' = a_\eta, \quad a_3' = a_\zeta; \qquad . \quad . \quad . \quad (8)$$

while the solution of (6) gives

$$\frac{x_1'}{(x\eta\zeta)} = \frac{x_2'}{(\xi x \zeta)} = \frac{x_3'}{(\xi\eta x)} = \frac{1}{(\xi\eta\zeta)}, \qquad . \quad . \quad (9)$$

so that the symbols α, cogredient with x, satisfy analogous equations

$$\frac{\alpha_1'}{(\alpha\eta\zeta)} = \frac{\alpha_2'}{(\xi\alpha\zeta)} = \frac{\alpha_3'}{(\xi\eta\alpha)} = \frac{1}{(\xi\eta\zeta)} \qquad (10)$$

Equations (8) hold for every symbol $a, b, c \ldots$, and (10) for every $\alpha, \beta, \gamma \ldots$. They also typify the general case when n columns $\xi, \eta, \ldots, \omega$ occur in (6) and n-rowed determinants in the denominators of (10). We can now prove the Fundamental Theorem.

7. Fundamental Theorem for the General Multilinear Form.

Every rational integral invariant of multilinear ground forms whose symbols are a, b, c $\ldots$, $\alpha, \beta, \gamma \ldots$ *is expressible as symbolic polynomials, the factors of whose terms are of three types*

$$a_\alpha, \quad (abc \ldots m), \quad (\alpha\beta\gamma \ldots \mu),$$

together with numerical coefficients.

The proof follows the lines of previous simpler cases, but requires two more preliminary lemmas which for clearness will be explained for the ternary case.

In the preceding formulæ (7), (8), (10), the coefficients ξ, η, ζ appear along with symbols $\alpha, \beta, \gamma \ldots$. Let us now consider a_ξ, a_η, a_ζ all to be of the same type as a_α, and $(\alpha\beta\xi)$, $(\alpha\xi\eta)$, $(\alpha\eta\zeta)$, &c., all to be of the same type as $(\alpha\beta\gamma)$. So we provisionally use inclusive symbols, θ, ϕ, ψ which may take the values

$$\theta, \phi, \psi = \xi, \eta, \zeta, \alpha, \beta \ldots .$$

LEMMA I.—*The effect of the Cayley operator* $\Omega \equiv \left| \frac{\partial}{\partial \xi_1} \; \frac{\partial}{\partial \eta_2} \; \frac{\partial}{\partial \zeta_3} \right.$ *upon a product of factors of types* a_θ, (abc), $(\theta\phi\psi)$, *which includes among its symbols one* ξ, *one* η, *and one* ζ, *is another product of the same types, but excluding* ξ, η, *and* ζ.

Such a product is typified by the following cases:

(i) $Q = (\xi\eta\zeta)N$,

(ii) $Q = (\xi\eta\theta)(\zeta\phi\psi)N$.

(iii) $Q = (\xi\eta\theta)\, a_\zeta N$,

(iv) $Q = (\xi\theta\phi)(\eta\theta'\phi')(\zeta\theta''\phi'')N$,

(v) $Q = (\xi\theta\phi)(\eta\theta'\phi')\, a_\zeta N$,

(vi) $Q = (\xi\theta\phi)\, a_\eta b_\zeta N$,

(vii) $Q = a_\xi b_\eta c_\zeta N$.

Here N denotes factors not containing ξ, η, ζ; and variations depending merely on derangement of the order ξ, η, ζ are excluded. Now apply the six-termed determinantal permutation $\dot{\xi}$, $\dot{\eta}$, $\dot{\zeta}$ to each of these cases. Then by the fundamental identity II, p. 48 (cf. §8, p. 41), the result in each of the cases (i) to (vi) is a derangement of the symbols ξ, η, ζ, θ and possibly ϕ, yielding $(\xi\eta\zeta)$ as factor. For example, in (vi) it gives

$$(\xi\eta\zeta)\,(a_\theta b_\phi - a_\phi b_\theta)N,$$

while in (vii) the result is $(\xi\eta\zeta)\,(abc)N$. Hence in all cases the process leaves the types unaffected and produces a factor $(\xi\eta\zeta)$. But by §5, p. 119 this operation is equivalent to

$$\left(\xi\eta\zeta \,\middle|\, \frac{\partial}{\partial\xi}\ \frac{\partial}{\partial\eta}\ \frac{\partial}{\partial\zeta}\right) Q = (\xi\eta\zeta)\,\Omega Q.$$

Hence dividing throughout by $(\xi\eta\zeta)$ the lemma is proved.

LEMMA II.—*The effect of the operator Ω on such a product involving several symbols ξ, η, ζ is a sum of such products of the same type but involving one ξ fewer, one η fewer, and one ζ fewer.*

Since the operator Ω is linear in $\frac{\partial}{\partial\xi_1}, \frac{\partial}{\partial\xi_2}, \frac{\partial}{\partial\xi_3}$ it operates on a function involving m factors, each linear in ξ, precisely as the ordinary rule for differentiating a product of m functions of one variable (cf. (33), p. 122)

$$\frac{d}{dx}(XY) = \frac{dX}{dx}Y + X\frac{dY}{dx}.$$

In fact it gives m terms, in each of which only one factor involving ξ undergoes operation. Likewise for η and ζ. Thus each such term behaves as in Lemma I, so that the second lemma is proved.

Proof of Fundamental Theorem.—

Let $I(A)$ denote a polynomial invariant in several sets of coefficients, typified by $A_{ijk\ldots}^{rst\ldots}$. Then if accents denote the corresponding new coefficients after transformation we have by hypothesis

$$I(A') = \phi(\xi_1, \ldots, \omega_n)I(A).$$

Although two contragredient transformations $x \rightarrow x'$, $u \rightarrow u'$ are now involved, they lead to one type of determinant $| M |$; and the argument already used in §2, p. 169, shows that $\phi(\xi_1, \ldots, \omega_n)$ can only be an integral power of $| M |$. This power may be zero or even negative.

For instance, the ternary bilinear form $a_x u_\alpha$ has an absolute invariant a_α, since

$$a'_{\alpha'} = \Sigma\, a_i' \alpha_i' = \frac{a_\xi(\alpha\eta\zeta) + a_\eta(\xi\alpha\zeta) + a_\zeta(\xi\eta\alpha)}{(\xi\eta\zeta)}, \text{ by (8) and (10),}$$
$$= a_\alpha(\xi\eta\zeta) \,/\, (\xi\eta\zeta) = a_\alpha,$$

after using the fundamental identity. Here the index of $(\xi\eta\zeta)$ in $\phi(\xi_1, \ldots, \omega_n)$ is zero.

We therefore take in the ternary case $I(A') = (\xi\eta\zeta)^w I(A)$ where w is an integer, positive, zero, or negative. Let this be thrown into symbolic form,

$$I(a_i', \ldots, b_i', \ldots, \alpha_r' \ldots, \ldots) = (\xi\eta\zeta)^w I(a_i, \ldots, b_j, \ldots, \alpha_r \ldots, \ldots).$$

Then by (8) and (10) the left-hand polynomial can be written

$$I\left(a_\xi, a_\eta, a_\zeta, b_\xi, \ldots, \frac{(\alpha\eta\zeta)}{(\xi\eta\zeta)}, \frac{(\xi\alpha\zeta)}{(\xi\eta\zeta)}, \frac{(\xi\eta\alpha)}{(\xi\eta\zeta)}, \ldots\right)$$

which is a polynomial in its arguments a_ξ, a_η, $\ldots$, $(\alpha\eta\zeta)$, $\ldots$ together with a common denominator, say $(\xi\eta\zeta)^\rho$, since the original function $I(A')$ is homogeneous originally in each set of coefficients and therefore finally in each set of symbols α' which it requires. Multiplying through by $(\xi\eta\zeta)^\rho$ we have, as our identity,

$$I(a_\xi, a_\eta, a\ , b_\xi, \ldots, (\alpha\eta\zeta), (\xi\alpha\zeta), (\xi\eta\alpha), \ldots)$$
$$= (\xi\eta\zeta)^{w+\rho} I(A).$$

The degree in ξ_1, ξ_2, ξ_3 is $w + \rho$ throughout; while ξ only enters the left member by way of factors of types

$$a_\xi, b_\xi, \ldots, (\xi\alpha\zeta), (\xi\eta\alpha) \ldots.$$

Since at least one such factor arises if merely a single actual coefficient A occurs in $I(A)$, it follows that $w + \rho$ must be a *positive* integer.

We now operate $w \div \rho$ times in succession with the Cayley operator

$$\Omega = \frac{\partial}{\partial \xi_1} \frac{\partial}{\partial \eta_2} \frac{\partial}{\partial \zeta_3}$$

and obtain, by Lemma II, an aggregate of the desired type on the left, with a non-zero numerical multiple of $I(A)$ on the right. For the process removes one ξ, η, and ζ at each stage, thereby freeing the left side entirely of ξ, η, ζ, because it does so on the right, at the same time preserving the desired types of factors on the left. Since exactly the same methods hold for the general as for this ternary case, this proves the theorem.

8. Covariants, Contravariants, and Mixed Concomitants.

Historically the functions which satisfy the typical invariant property

$$\psi(A') = \phi(\xi_1, \dots, \omega_n)\, \psi(A) \quad . \quad . \quad . \quad (11)$$

have been classified in the following manner:

Concomitants

- (i) *Invariants,*
- (ii) *Covariants,*
- (iii) *Contravariants,*
- (iv) *Mixed concomitants.*

Functions (i) involve coefficients of ground forms only: (ii) involve variables $x_1, \dots, x_n$ besides; (iii) involve the contragredient variables $u_1, \dots, u_n$ instead of x; (iv) involve any variables that may exist. Thus if

$$f_1 = a_x^2, \quad f_2 = b_x^2, \quad f_3 = c_x^2,$$

are three ternary quadratics, we have as instances of these four types of concomitant

$$\begin{array}{rl} \text{(i)} & (abc)^2 \\ \text{(ii)} & (abc)\,a_x b_x c_x \\ \text{(iii)} & (abu)\,(bcu)\,(cau) \\ \text{(iv)} & (abu)\,a_x b_x \\ \text{(iv}a\text{)} & u_x \end{array} \qquad (12)$$

For by the general fundamental theorem, concomitants are composed of symbolic factors a_α, (abc), $(\alpha\beta\gamma)$, while by (5), p. 202, u

behaves like symbols a, b, c and x like α, β, γ. Hence a_x, (abu), (abc) are possible factors of a concomitant. Also exactly two a's, two b's, two c's go to form expressions (12), so that they are actual concomitants of the quadratics.

Here (ii) is the cubic covariant of three ternary quadratics which we have already seen to be their Jacobian (§1, p. 182): (iii) is a cubic contravariant, because on expansion it is a cubic in u_1, u_2, u_3: (iv) is a mixed form of orders (1, 2), while (iv$a$) is sometimes called the *absolute concomitant* of the field.

This classification, however, does not go very deep, because all concomitants can be treated as invariants, merely by adjoining suitable ground forms. Since, when $x \rightarrow x'$, $u \rightarrow u'$,

$$u_1 x_1 + u_2 x_2 + \ldots + u_n x_n = u_1' x_1' + \ldots + u_n' x_n',$$

we may regard the variables $u_1, \ldots, u_n$ as coefficients of a certain linear form a_x, where $u_i = a_i$. Correlatively the variables x may be regarded as coefficients of a certain linear form u_α, where $x_i = \alpha_i$. And, in fact, *the problem of finding covariants of a ground form* f *is indistinguishable from that of finding invariants of two ground forms*—f *and a linear form* u_x, *treating* u *as the variable in the latter* (cf. p. 145).

If we treat $u, v, w, \ldots, x, y, z, \ldots$ as such coefficient sets, and use U, X for variables, we reduce all concomitants involving ground forms f together with $u, v, w, \ldots, x, y, z, \ldots$ to invariants of forms f together with linear forms $u_X, v_X, w_X, \ldots, U_x, U_y, U_z, \ldots$.

Hence the fundamental theorem covers the case of all concomitants of all four types (i), (ii), (iii), and (iv).

Example.—

The polar process $\sum_\mathrm{i} \mathrm{y_i} \dfrac{\partial}{\partial \mathrm{x_i}}$ *is a particular case of the Aronhold process* (p. 140). For it is the latter applied to a linear form U_x. The same remark is true of any polar process $\left(v\dfrac{\partial}{\partial u}\right)$, &c. Hence polarization is an *invariant process* (p. 141).

9. **Convolution and Resolution.**

The absolute concomitant (iv$a$) can take various forms since by the methods of §8, p. 86, x can be resolved into $n-1$ components $v, w, \ldots, t$ or correlatively u can be resolved into $n-1$ such as $y, \ldots, z$.

For ternary forms if x, y, z are three such points, and $u = \overline{yz}$, $v = \overline{zx}$, $w = \overline{xy}$ are three lines forming the sides of their triangle, then the expressions

$$(xyz) = u_x, \quad (uvw) = (yz \,.\, zx \,.\, xy) = (xyz)^2$$

are absolute concomitants. Similarly for higher fields. This process of replacing one by many variables is called *resolution*, the converse being *composition* or *convolution*.

Exactly the same processes apply to the symbols; one α can be replaced by $n-1$ such as $a', a'', \dots, a^{(n-1)}$; and one a by $n-1$ such as $\alpha', \alpha'', \dots, \alpha^{(n-1)}$.

Example.—

The ternary quadric $a_x{}^2$ could be symbolized by $(\alpha' \alpha'' x)^2$, where $a_1 = (\alpha' \alpha'')_{23}$, &c.

10. The Fundamental Theorem for the General Case.

In order to cover all cases contemplated in the classification of (12) above, we must consider a possible ground form with several sets of variables and perhaps distinct fields of linear transformation. The case of absolutely distinct fields will be dealt with later in §4, p. 240, but to round off the present investigation we must contemplate the ground form which contains rth compound co-ordinate sets, $r = 2, 3, \dots, n-2$, the case $r = n-1$ having already been included.

As an example, consider the multilinear form in three variables x, y, z

$$f = a_x b_y c_z = \Sigma A_{ijk} x_i y_j z_k.$$

Suppose that it is not the most general polynomial in y and z. but changes sign when y and z are interchanged: namely the coefficients A_{ijk} obey the law

$$A_{ijk} = -A_{ikj}$$

for all values of j and k. Then after interchange we write

$$-f = a_x b_z c_y = \Sigma A_{ijk} x_i z_j y_k,$$

whence by subtraction

$$\begin{aligned} 2f &= a_x(b_y c_z - b_z c_y) = \Sigma A_{ijk} x_i (y_j z_k - y_k z_j) \\ &= a_x(bc \mid yz) = \Sigma A_{ijk} x_i (yz)_{jk}. \end{aligned}$$

This is now a bilinear form in two sets, x and a second compound $\pi_2 = \overline{yz}$. We note that f is symbolized by two symbolic linear factors, one for x as before, and the other, $(bc \mid yz)$, a type already familiar through the theorem of corresponding matrices.

Furthermore, if

$$jk,\ l \ldots m$$

are algebraic complements among n suffixes $jkl \ldots m$ deranged from $123 \ldots n$, and if variables $v \ldots w$, contragredient to y, are introduced, we have by the principle of duality

$$(yz)_{jk} = (v \ldots w)_{l \ldots m}, \qquad (bc \mid yz) = (bcv \ldots w). \qquad (13)$$

But this last is an ordinary bracket factor of n symbols, which we can also write as

$$(bcv \ldots w) = (bcp_{n-2}), \quad . \quad . \quad . \quad . \quad (14)$$

where p_{n-2} is the n-rowed matrix of currency $(n - 2)$.

More generally, by the same reasoning, a form

$$g = b_y c_z d_{y'} e_{z'} \ldots$$

alternating in y, z, also in y', z', also in y'', z'', and so on for q pairs, is symbolized more explicitly by

$$(bc \mid yz)(de \mid y'z') \ldots = (bc \mid \pi_2)(de \mid \pi_2') \ldots = (bc\, p_{n-2})(de\, p'_{n-2}) \ldots,$$

and if these second compounds $\pi_2, \pi_2' \ldots$ happen to be equal, the form is

$$(bc \mid \pi_2)(de \mid \pi_2) \ldots$$

of degree q in π_2.

What has been said of these second compounds would also apply to other values of r. In this way we are led to consider such a ground form as

$$F \equiv a_x a_y' a_z'' \ldots (bc \mid \pi_2)(b'c' \mid \pi_2') \ldots (def \mid \pi_3)(d'e'f' \mid \pi_3') \ldots, \quad (15)$$

with q_1 factors like a_x, q_2 involving second, q_3 third, ..., and finally q_{n-1}, $(n - 1)$th compounds of type u. When $x = y = z$, $\pi_2 = \pi_2'$, ..., $\pi_r = \pi_r'$, ..., this is called a mixed ground form of orders

$$[q_1, q_2, \ldots, q_{n-1}], \qquad q_r \geqslant 0, \quad . \quad . \quad . \quad (16)$$

in the variables $x, \pi_2, \dots, \pi_{n-1} = u$. At a later stage it will be seen, as Clebsch was the first to prove, that this is the most general type of ground form necessarily occurring in the invariant theory of one set of variables and all its compounds. At present we have merely found that such a form *may* arise. With this theorem of Clebsch ultimately in view to give significance to such a form, once more we adapt the Fundamental Theorem to include F as a possible ground form.

11. Proof of the Fundamental Theorem.

To effect this, it is only necessary to show that the new type of factor in F, due to the rth compound ($r = 2, 3, \dots, n-2$) leads to the same general final result as before, namely an aggregate of types

$$a_\alpha, \quad (abc \dots m), \quad (\alpha\beta\gamma \dots \mu).$$

The typical new type of factor in F is

$$(A_r \mid \pi_r) \equiv (a_1 a_2 \dots a_r \mid \pi_r) \equiv (a_1 a_2 \dots a_r \mid yz \dots t),$$
$$(r = 2, 3, \dots, n-2),$$

where each a_i is a symbol cogredient with a or u, and where $y, z, \dots, t$ are r variables cogredient with x. We denote the expanded form of this type of factor by

$$\Omega_{A_r} a_{1y} a_{2z} \dots a_{rt},$$

where Ω_{A_r} is the determinantal permutation of $r!$ terms previously denoted by

$$\dot{a}_1 \dot{a}_2 \dots \dot{a}_r.$$

Let this be done for each such factor, where $r = 2, 3, \dots, n-2$, so that F is now expressed symbolically in terms of symbols and variables a, u, α, x as before, except that certain groups of symbols a undergo determinantal permutation.

For instance, if A denotes aa' and B, bb', the form

$$F = (A \mid yz)(B \mid xt) = (aa' \mid yz)(bb' \mid xt)$$

is now written

$$F = \Omega_A a_y a_z' \Omega_B b_x b_t' = \Omega_A \Omega_B a_y a_z' b_x b_t'.$$

It consists of four terms due to simultaneous derangement of aa' and of bb'. Clearly we can write it in many ways

$$F = \Omega_A \Omega_B a_y a_z' b_x b_t' = \Omega_B \Omega_A a_y \ldots = -\Omega_A \Omega_B a_y' a_z b_x b_t',$$

and so on.

Each operating symbol Ω now denotes a group implicitly convolved (§10, p. 46) in the operand, as shown by the suffix of Ω.

If there are ν such groups $A, B, \ldots, K$ in the form F we denote the expanded form of F by

$$F' = \Omega_A \Omega_B \ldots \Omega_K a_{1y} \ldots k_{r_\nu z} G,$$

where G denotes all other factors of F not so affected.

Now any actual coefficient $A'^{st\ldots}_{ij\ldots}$ of F after linear transformation is obtained symbolically by substituting one or other of $\xi, \eta, \ldots, \omega$ for each variable $y \ldots$ occurring in F. Hence it only differs from the result in the previous case by having ν groups of symbols implicitly convolved.

Since these symbols $a_1, \ldots, k_{r_\nu}$ do not include $\xi, \eta, \ldots, \omega$, the Corollary IV, §11, p. 49, applies when we permute $\xi, \eta, \ldots, \omega$, as is the case when the Cayley operator Ω acts upon a supposed invariant. Thus we can perform each step of the previous proof of the Fundamental Theorem, and obtain the same result as before, only modified in this respect, that

In the final aggregate of types

$$a_\alpha, \quad (abc \ldots m), \quad (\alpha\beta\gamma \ldots \mu)$$

every group of symbols $\mathrm{a}_1, \mathrm{a}_2, \ldots, \mathrm{a}_\mathrm{r}$ *which were convolved in the original ground form, still preserve this property implicitly in the symbolic form of the invariant.*

With this proviso the theorem has now been completely established for mixed ground forms in any number of compound variables for the field of order n.

Example.—

In quaternary forms three types of variable are used

$$\overline{uvw} = x, \quad \overline{uv} = p = \overline{xy}, \quad u = \overline{xyz}.$$

The form

$$F = \Sigma k_{ij} p_{kl} = (aa'p) = (aa'uv)$$

is called a linear complex. If equivalent symbols are used, we write

$$(aa'uv) = (bb'uv) = (cc'uv) = (dd'uv).$$

Hence a possible invariant of weight two has the term $(aa'bc)(b'c'dd')$, which gives rise to a four-term series due to the couple of alternatives b, b'; $-b', b$ and of c, c'; $-c', c$. According to the Fundamental Theorem this series would be an invariant,

$$I = \Omega_B \Omega_C (aa'bc)(b'c'dd'),$$

where Ω_B denotes the derangement of bb', and Ω_C of cc'. Such a notation is preferable to the dot notation when several independent derangements proceed simultaneously.

EXAMPLES[1]

1. If the above invariant is written as

$$-I = \Omega_B \Omega_C (b'aa'c)(c'dd'b) = \{BACD\},$$

where A denotes aa', and B, bb', and so on, prove that

$$\{BACD\} = \{CDBA\} = \{BDCA\} = \{CABD\}.$$

2. Using identity I, p. 45, to convolve cc' explicitly in I, prove that

$$\{BACD\} + \{BCAD\} = 2(AC)(DB),$$

where $(AC) = (aa'cc')$. Hence prove that

$$I = (AB)(CD) + (AC)(BD) - (AD)(BC).$$

This procedure renders each of the four convolutions A, B, C, D *explicit.* Originally two were implicit.

3. Let $J = \Omega_C (abb'c)(c'dd'd'') = (aBC\delta)$, say. Then J is a simultaneous invariant of a plane a, two linear complexes B and C, and a point δ. Here a is cogredient with u, B and C each with p, and δ with x. Prove the identity

$$(aBC\delta) + (aCB\delta) + (BC)a_\delta = 0.$$

4. An invariant exists, linear in the coefficients of three linear complexes and two planes.

Let $K = \Omega_C (abb'c)(c'dd'e) = (aBCDe)$, where B, C, D refer to three complexes, and a, e to two planes. Prove that $(aBCDe) = -(eDCBa)$, and that

$$(aBCDe) + (aCBDe) + (BC)(aDe) = 0.$$

5. Prove that $(uBCDu)$, which is a special case of the expression K, is an alternating function of B, C, D. Thus $(uBCDu) = -(uCBDu) = (uCDBu) = \&c.$

6. Discuss the corresponding dual invariant

$$K' = \Omega_B \Omega_D (\alpha b)(b'cc'd)(d'\varepsilon)$$

of two points α, δ and three complexes B, C, D.

---

[1] For references see §13, p. 330.

CHAPTER XIII

Symbolic Methods of Reduction

1. The Fundamental Identities.

The First Fundamental Theorem has provided a uniform in which any concomitant can be dressed, together with a means of constructing as many as we like. The next stage in the theory is to learn how such symbolic forms behave, to simplify or reduce or transform them. Like ordinary numbers and like matrices, they have their fundamental properties or rules of combination. For binary forms these run as follows:

(i) $(ab) = -(ba)$,

(ii) $(bc)a_x + (ca)b_x + (ab)c_x = 0$,

(iii) $(bc)(ad) + (ca)(bd) + (ab)(cd) = 0$,

(iv) $a_x b_y - a_y b_x = (ab)(xy)$.

(v) The interchange of equivalent symbols.

All these properties have been established and illustrated. It is perhaps the last which presents the greatest novelties and apparent drawbacks of the symbolic methods, because it leads to alternative expressions for one and the same invariant, without suggesting which of them is to be taken as the simpler, Thus these three expressions

$$(ab)(ac)b_x c_x, \quad \tfrac{1}{2}(ab)^2 c_x^2, \quad (ba)(bc)c_x a_x$$

are all equal to $(a_0 a_2 - a_1^2) \times f$ for the binary quadratic

$$f = a_0 x_1^2 + 2a_1 x_1 x_2 + a_2 x_2^2 \ = \ a_x^2 = b_x^2 = c_x^2.$$

Again, as in the theory of determinants, very substantial functions such as $(bc)(ca)(ab)$ can vanish identically—but for a new reason, owing to the operation (v). For if a, b are equivalent, then on interchange

$$(bc)\,(ca)\,(ab) = (ac)\,(cb)\,(ba),$$

and by three applications of (i), this last is $-(ca)\,(bc)\,(ab)$ whence $2(bc)\,(ca)\,(ab) = 0$ or $(bc)\,(ca)\,(ab) = 0$. Cf. §3, p. 18.

For forms in n variables analogous properties hold, but owing to the two types of symbol a and α with contragredient behaviour, there are now the following laws:

(i) $(ab\ldots m) = -(ba\ldots m) = \ldots,$
$\quad(\alpha\beta\ldots\mu) = -(\beta\alpha\ldots\mu) = \ldots,$

(ii) $\Pi_1 = (ab\ldots m)\,(n\alpha)$
$$\qquad - (nb\ldots m)\,(a\alpha) + \ldots + (-)^n\,(nab\ldots)\,(m\alpha) \equiv 0,$$
$\quad\Pi_1' = (\alpha\beta\ldots\mu)\,(\nu a)$
$$\qquad - (\nu\beta\ldots\mu)\,(\alpha a) + \ldots + (-)^n\,(\nu\alpha\beta\ldots)\,(\mu a) \equiv 0,$$

(iii) $\Pi_2 = (ab\ldots m)\,(uv\ldots w) - (ub\ldots m)\,(av\ldots w) + \ldots \equiv 0,$
$\quad\Pi_2' = (\alpha\beta\ldots\mu)\,(xy\ldots z) - (x\beta\ldots\mu)\,(\alpha y\ldots z) + \ldots \equiv 0,$

(iv) $\Pi_3 = (ab\ldots m)\,(\alpha\beta\ldots\mu) - \Sigma \pm (a\alpha)\ldots(m\mu) \equiv 0,$

(v) Interchange of equivalent symbols a with b or α with β.

Here Π_1 and Π_1' are dual identities, and so also are Π_2 and Π_2'. They follow from (IV), §9, p. 46, while Π_3 is a statement of the product theorem of two determinants.

2. The Second Fundamental Theorem.

The question now arises, granted that these identities teach us something of the properties of symbolic forms, do they leave any gaps, do any properties escape? The answer is given by a very remarkable theorem usually called the *Second Fundamental Theorem*, which was originally proved for binary forms by Gordan.[1] Later a proof when $n = 3$ was given by E. Study[2] and for the general case by E. Pascal,[3] and recently for compound co-ordinates and for restricted transformations by R. Weitzenböck.[4]

---

[1] Cf. Gordan, *Invariantentheorie*, Bd. II, §117: Pascal, *Battaglini*, **26** (1888). Grace and Young, *Algebra of Invariants* (1904), p. 368.

[2] Study, *Methoden* . . . (Leipzig, 1889), p. 75 and p. 204.

[3] Pascal, *Mem. del R. Acc. dei Lincei*, V, **4a** (1888).

[4] *Wiener Berichte*, **122** (1913). Cf. *Invariantentheorie*, pp. 98–113, by the same author.

The theorem states:

Every identity satisfied by invariants (concomitants) can ultimately be expressed by the fundamental identities together with the principle of the interchange of equivalent symbols and the laws of ordinary algebra for the combination of the three elementary types a_α, $(abc \dots m)$, $(\alpha\beta\gamma \dots \mu)$.

In other words the symbolic theory is a complete and self-contained discipline; and from the logical point of view it is this which gives it a permanent importance.

The theorem tells us that any such polynomial identity Π symbolizing such an actual identity in coefficients and variables of ground forms can be expressed as

$$\Pi = A_1\Pi_1 + A_2\Pi_1' + A_3\Pi_2 + A_4\Pi_2' + A_5\Pi_3 \equiv 0,$$

where, of course, these coefficients A_i need not vanish identically.

3. Binary Quadratic Forms. Reducibility.

As an illustration of the symbolic principles of reduction, we shall consider the problem of finding all possible different types of concomitants of a set of binary quadratics.

Let

$$\begin{aligned} f_1 &= a_{11}x_1^2 + 2a_{12}x_1x_2 + a_{22}x_2^2 = a_x^2, \\ f_2 &= b_{11}x_1^2 + 2b_{12}x_2 + b_{22}x_2^2 = b_x^2, \end{aligned} \qquad \dots (1)$$

$f_3 = c_x^2$, $f_4 = d_x^2$, &c., so that for each letter $a, b, c, d \dots$ we have relations

$$a_{ij} = a_{ji} = a_i a_j = a_j a_i, \quad i, j, = 1, 2. \qquad \dots (2)$$

If we consider merely one variable x, together with these coefficients, every concomitant is expressible, by the fundamental theorem, as an aggregate of factors (ab), a_x where every letter a or b is duplicated in each term. For example,

$$V = a_x(ab)b_x \qquad W = c_x(cd)d_x$$

are covariants, and so also is

$$(ab)\ (cd)a_x b_x c_x d_x,$$

which is easily seen to be the same as VW, and is said to be *reducible* because it is expressible in terms of simpler covariants. But V itself is not reducible by such factorizing because no partition into factors a_x, (ab), b_x gives an actual covariant in non-symbolic form. The covariants V and W are said to be of the same type because they merely differ by choice of symbols a, b, c, d of the ground forms; their structure is the same.

Rejecting forms of the same type and obviously reducible forms, we can immediately write down a list of possible single-term invariants. They would be

$$(ab)^2 = D_{12},$$
$$(ab)\,(bc)\,(ca) = D_{123},$$
$$(ab)\,(bc)\,(cd)\,(da) = D_{1234},$$

with a similar list of covariants

$$a_x{}^2 = C_1 = f_1,$$
$$a_x\,(ab)\,b_x = C_{12},$$
$$a_x\,(ab)\,(bc)\,c_x = C_{123},$$
$$a_x\,(ab)\,(bc)\,(cd)\,d_x = C_{1234},$$

For these are the only possible structures, involving two a's, two b's . . . which follow the conditions laid down. Thus if $C_{ij\ldots}$, $D_{ij\ldots}$ refer to quadratics f_i, f_j . . . , every concomitant must be a polynomial in $C_{ij\ldots}$, $D_{ij\ldots}$. We may, however, reject all but the first two entries in each list as expressible in terms of these simpler concomitants.

In fact by a fundamental identity since

$$(bc)\,a_x + (ab)\,c_x = (ac)\,b_x,$$

therefore, squaring both sides,

$$(bc)^2 a_x{}^2 + (ab)^2 c_x{}^2 + 2a_x\,(ab)\,(bc)\,c_x = (ac)^2 b_x{}^2,$$

or

$$2C_{123} = D_{13}f_2 - D_{12}f_3 - D_{23}f_1.$$

On dividing by 2 this expresses C_{123} in terms of the simpler D_{ij}, f_i. Likewise for C_{ijk}.

Next by polarizing this identity with regard to y_1, y_2, and remembering $\left(y\dfrac{\partial}{\partial x}\right)a_x c_x = a_y\,c_x + a_x\,c_y$, we have

$$a_y\,(ab)\,(bc)c_x + a_x(ab)\,(bc)c_y = D_{13}\,b_x\,b_y - D_{12}\,c_x\,c_y - D_{23}\,a_x\,a_y$$

identically true for all values of y.

In particular, if $y_2 = -d_1 d_x$, $y_1 = d_2 d_x$, then

$$c_y = (cd)\, d_x;$$

whence

$$a_x (ab)(bc)(cd) d_x + d_x (ad)(ab)(bc) c_x = D_{13} b_x (bd) d_x - \&c.$$

But, by a fundamental identity,

$$\begin{aligned} a_x (ab)(bc)(cd) d_x - d_x (ab)(bc)(ad) c_x &= d_x (ab)(bc)(ca) d_x \\ &= D_{123} f_4. \end{aligned}$$

Adding these two results,

$$2a_x (ab)(bc)(cd) d_x = D_{13}\, C_{24} - \ldots + D_{123} f_4,$$

which expresses C_{1234} polynomially in terms of D_{123}, D_{ij}, C_{ij}, f_i. Similarly by putting $y_2 = -d_1 (de) e_x$, &c., we reduce C_{12345}; and so on for all successive entries in the column of C's.

Further by squaring the identity

$$(bc)(ad) + (ab)(cd) = (ac)(bd)$$

we reduce D_{1234}, or what is the same thing by putting $x_1 = d_2$, $x_2 = -d_1$ in C_{123} we effect this reduction. Each further entry of the D column is reduced by similar substitution for x. Thus we have proved the theorem:

All concomitants of any number of binary quadratics are expressible in terms of four types:

$$a_x^2, \quad (ab) a_x b_x, \quad (ab)^2, \quad (bc)(ca)(ab).$$

Corollary.—*Since* $(ab) a_x b_x = -(ba) a_x b_x$, *it vanishes identically when the two quadratics are the same, by interchanging equivalent symbols. Similarly for* (bc) (ca) (ab). Cf. §1, p. 214.

Hence $C_{ii} = 0$, $D_{iij} = 0$, leaving f_i, C_{ij}, D_{ii}, D_{ij}, D_{ijk} (where i, j, k are unequal) as the only possible irreducible concomitants of the quadratics. All these, except the last, have already been discussed in Chapter VIII.

We now see that they are completely typical of all possible concomitants. By no possibility can any of them be expressed rationally and integrally in terms of the others, as the reader will see if an attempt is made to do so. Accordingly they are said to form a complete system, and every polynomial concomitant is expressible as a polynomial function of members of the complete system

$$\phi(f_i,\ C_{ij},\ D_{ii},\ D_{ij},\ D_{ijk}).$$

The following table can now be made.

COMPLETE SYSTEM OF BINARY QUADRATICS

Number of Ground Forms.	Invariants.		Covariants.	
	Degree 2.	Degree 3.	Degree 1.	Degree 2.
1	1	0	1	0
2	3	0	2	1
n	$\frac{n(n+1)}{2}$	$\frac{n(n-1)(n-2)}{6}$	n	$\frac{n(n-1)}{2}$

4. Significance of the Complete System.

In §9, p. 139, we found the discriminant of the pencil of quadratics $U + \lambda V$. Hence the discriminant for $f_1 + \lambda f_2$ will be

$$(a_{11}a_{22} - a_{12}^2) + \lambda(a_{11}b_{22} + a_{22}b_{11} - 2a_{12}b_{12}) + \lambda^2(b_{11}b_{22} - b_{12}^2).$$

Symbolically this is

$$\tfrac{1}{2}\{(aa')^2 + 2\lambda(ab)^2 + \lambda^2(bb')^2\},$$

where $f_1 = a_x^2 = a'^2_x$, $f_2 = b_x^2 = b'^2_x$. Thus the discriminant of the pencil $f_1 + \lambda f_2$ is

$$\tfrac{1}{2}(D_{11} + 2\lambda D_{12} + \lambda^2 D_{22}).$$

Hence D_{ii} is the discriminant of the quadratic f_i, and D_{ij} is the simultaneous invariant, sometimes called the harmonic invariant of two quadratics f_i and f_j.

Next, the other type of invariant can be written as,

$$D_{123} = (bc)(ca)(ab) = -\begin{vmatrix} a_1^2 & a_1a_2 & a_2^2 \\ b_1^2 & b_1b_2 & b_2^2 \\ c_1^2 & c_1c_2 & c_2^2 \end{vmatrix}$$

which is easy to verify. Hence in terms of the actual coefficients it is

$$\begin{array}{ccc} a_{11} & a_{12} & a_{22} \\ b_{11} & b_{12} & \\ c_{11} & c_{12} & c_{22} \end{array}$$

This important invariant is called the determinant of the coefficients of three quadratics. When it vanishes the quadratics are said to be in involution. But such a determinant vanishes when and only when the three quadratics are linearly dependent, §2, p. 75. In fact values λ_1, λ_2, λ_3 exist such that

$$\lambda_1 f_1 + \lambda_2 f_2 + \lambda_3 f_3 = 0$$

identically for all values of x_1, x_2.

Since a binary quadratic has three coefficients, three quadratics naturally lead to a square coefficient matrix. If the rank of this is less than three, the determinant D_{123} vanishes.

Next, there is the covariant $C_{12} = (ab)\,a_x\,b_x$, which is the Jacobian, already noticed in §12, p. 144. For

$$\begin{vmatrix} \dfrac{\partial a_x^2}{\partial x_1} & \dfrac{\partial b_x^2}{\partial x_1} \\ \dfrac{\partial a_x^2}{\partial x_2} & \dfrac{\partial b_x^2}{\partial x_2} \end{vmatrix} = \begin{vmatrix} a_x a_1 & b_x b_1 \\ a_x a_2 & b_x b_2 \end{vmatrix} = (ab)\,a_x\,b_x.$$

5. Canonical Form of Two Binary Quadratics.

Suppose the Jacobian $J = (ab)\,a_x\,b_x$ has two distinct linear factors X, Y. Since these are linear in x_1, x_2 we can take them as new variables. Let the quadratics f_1, f_2 now be

$$A_{11}X^2 + 2A_{12}XY + A_{22}Y^2, \quad B_{11}X^2 + 2B_{12}XY + B_{22}Y^2. \quad (3)$$

Reference to (43), p. 144, shows that J takes the required form λXY if, and only if,

$$A_{11}B_{12} - A_{12}B_{11} = 0, \quad A_{12}B_{22} - A_{22}B_{12} = 0. \qquad (4)$$

For these vanishing expressions are the coefficients of X^2 and Y^2 in J. This requires either the original quadratics to differ by a mere constant multiplier so that

$$A_{11} : A_{12} : A_{22} = B_{11} : B_{12} : B_{22} \quad . \quad . \quad . \quad (5)$$

or else $A_{12} = B_{12} = 0$. In general the first alternative is not true; hence two distinct quadratics whose Jacobian has distinct linear factors can be expressed as the sums of squares

$$f_1 = A_{11}X^2 + A_{22}Y^2, \quad f_2 = B_{11}X^2 + B_{22}Y^2. \quad . \quad (6)$$

Further, if the quadratics have non-vanishing discriminants, none of A_{11}, A_{22}, B_{11}, B_{22} vanish, otherwise f_1 or f_2 becomes a perfect square. We can therefore take

$$X_1 = \surd A_{11}X, \quad X_2 = \surd A_{22}Y, \quad . \quad . \quad . \quad (7)$$

as a new linear transformation, finally obtaining

$$f_1 = X_1^2 + X_2^2, \quad f_2 = A_1X_1^2 + A_2X_2^2, \quad . \quad (8)$$

where $A_1 = B_{11}/A_{11}$, $A_2 = B_{22}/A_{22}$. This is called the canonical form of two binary quadratics, to which any such ground forms can be reduced provided that they satisfy the conditions incidentally utilized above.

If we write M for the modulus of the transformation from the original variables x_1, x_2 to X_1, X_2 then the Jacobian and discriminants of the canonical forms are (p. 144) multiples of the original Jacobian and discriminants, namely

$$MJ, \quad M^2D_{11}, \quad M^2D_{12}, \quad M^2D_{22}.$$

Thus by direct calculation we find

$$MJ = (A_2 - A_1)X_1X_2, \quad M^2D_{11} = 2, \quad M^2D_{12} = A_1 + A_2,$$
$$M^2D_{22} = 2A_1A_2. \quad . \quad . \quad . \quad . \quad . \quad . \quad (9)$$

Hence the concomitants satisfy an identical relation

$$2J^2 = 2D_{12}f_1f_2 - D_{11}f_2^2 - D_{22}f_1^2, \quad . \quad . \quad (10)$$

as is readily verified from the canonical forms. This is the only relation which exists between these six forms, for in (8) and (9) the five quantities M, A_1, A_2, X_1, X_2 are arbitrary and any further relation between the concomitants would reduce the number of these arbitrary parameters to four.

EXAMPLES

1. Find the discriminant of J, and its significance.

2. Prove symbolically that the simultaneous invariant of the Jacobian of two quadratics f_1, f_2 and another quadratic f_3, is D_{123}.

6. Extension to Forms of Higher Order.

Before giving a geometrical interpretation of these results, it is worth noting how they may be extended. Since the symbols a_1, a_2, b_1, b_2 . . . behave like ordinary numbers, any identity, or relation, established for quadratics necessarily gives information about cubics and higher forms.

EXAMPLES

1. Just as $(ab)a_x b_x$ is the Jacobian of two quadratics, $(ab)a_x^{m-1}b_x^{n-1}$ is, to a constant numerical factor mn, the Jacobian of an m-ic a_x^m and an n-ic b_x^n.

2. Since $2(ab)(ac)b_x c_x = (ab)^2 c_x^2 + (ac)^2 b_x^2 - (bc)^2 a_x^2$ (§3, p. 216), it follows that

$$2(ab)(ac)a_x^{m-2}b_x^{n-1}c_x^{p-1} = (ab)^2 a_x^{m-2}b_x^{n-2}c_x^{p} + (ac)^2 a_x^{m-2}b_x^{n}c_x^{p-2} - (bc)^2 a_x^{m}b_x^{n-2}c_x^{p-2},$$

when $$m > 1,\ n > 1,\ p > 1.$$

In such identities the significant parts of each term are the bracket factors, for, if these are known, the x-factors a_x, b_x, c_x follow automatically to make up the requisite number of symbols for each term.

7. Transvectants.

Definition of Transvectant.—*The covariant* $(\mathrm{ab})^{\mathrm{r}}\,\mathrm{a_x}^{\mathrm{m-r}}\,\mathrm{b_x}^{\mathrm{n-r}}$ *of binary quantics* $\mathrm{f} = \mathrm{a_x}^{\mathrm{m}}$, $\phi = \mathrm{b_x}^{\mathrm{n}}$ *is their* r*th transvectant, and is often written*

$$(f, \phi)^r.$$

If r exceeds the lesser of m and n, the transvectant is zero: if $r = m = n$, it is an invariant: and if $r = 1$ it is, to a constant factor, their Jacobian, written (f, ϕ) with the index omitted.

We may easily prove that all odd transvectants of a form f with itself vanish identically, and all even transvectants give its covariants of degree two. For if

$$f = a_x^m = b_x^m,$$

then

$$(f,f)^r = (ab)^r a_x^{m-r} b_x^{m-r} = (-)^r (ba)^r a_x^{m-r} b_x^{m-}$$

which leads to a zero result after interchanging equivalent symbols, if r is odd. Also by the fundamental theorem the only covariants of degree two are polynomials in (ab), a_x, b_x: hence they are transvectants.

Another important case is the Hessian of a binary m-ic, namely,

$$\begin{matrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{matrix} = m^2(m-1)^2 \begin{matrix} a_x^{m-2} a_1^2 & b_x^{m-2} b_1 b_2 \\ {}^{m-2} a_1 a_2 & b_x^{m-2} b_2^{2} \end{matrix}$$

$$= m^2(m-1)^2 \begin{matrix} a_1 & b_1 \end{matrix} \; a_1 b_2 \, a_x^{m-2} \, b_x^{m-2}$$

$$= \tfrac{1}{2} m^2(m-1)^2 (ab)^2 a_x^{m-2} b_x^{m-2}$$

$$= \tfrac{1}{2} m^2(m-1)^2 (f, f)^2 = \tfrac{1}{2} m^2(m-1)^2 H.$$

Here we have started with the Jacobian of the two first polars $\dfrac{\partial f}{\partial x_1}$, $\dfrac{\partial f}{\partial x_2}$ of a binary m-ic, and have thrown it into symbolic form, thereby requiring two equivalent symbols because the original determinant was of degree two in the coefficients of f. The result $(f, f)^2$ shows that the Hessian is a covariant, a feature which can easily be generalized for the Hessian of n variables.

EXAMPLES

1. The Hessian does not exist for a linear form.

2. The Hessian of a quadratic is its discriminant.

3. The Hessian of a binary cubic is a quadratic.

4. The Hessian of a binary quartic is a quartic.

5. The coefficient of the highest power of x_1 in the Hessian of $(a_0, a_1, \ldots, a_n \between x_1, x_2)^n$ is $a_0 a_2 - a_1^2$, to a constant numerical factor.

6. Write in full the Hessian $\left| \dfrac{\partial^2 f}{\partial x_i \partial x_j} \right|$ of a ternary form

$$f = a_x^m = b_x^m = c_x^m,$$

and show that neglecting a numerical factor its symbolic form is

$$(abc)^2 a_x^{m-2} b_x^{m-2} c_x^{m-2}.$$

7. The bracket factors of the Hessian of an n-ary form $a_x^p = b_x^p$ = &c., are $(abc \ldots m)^2$. The Hessian is of weight two and degree n.

8. Reducibility of Jacobians.

We can now prove two theorems concerning binary Jacobians.

THEOREM I.—*The Jacobian of two quantics, one of which is a Jacobian, is reducible.*

This is often quoted as: *the Jacobian of a Jacobian is reducible*, and it is true of forms of general order n.[1]

Proof.—

Let (f, ϕ) denote the Jacobian $(ab)a_x^{m-1}b_x^{n-1}$ of $f = a_x^m$ and $\phi = b_x^n$. Then the Jacobian of (f, ϕ) and $\psi = c_x^p$ is

$$J = \begin{matrix} \dfrac{\partial(f, \phi)}{\partial x_1} & \dfrac{\partial \psi}{\partial x_1} \\[2ex] \dfrac{\partial(f, \phi)}{\partial x_2} & \dfrac{\partial \psi}{\partial x_2} \end{matrix}$$

$$\begin{matrix} (m-1)(ab)a_x^{m-2}b_x^{n-1}a_1 + (n-1)(ab)a_x^{m-1}b_x^{n-2}b_1, & pc_x^{p-1}c_1 \\ (m-1)(ab)a_x^{m-2}b_x^{n-1}a_2 + (n-1)(ab)a_x^{m-1}b_x^{n-2}b_2, & pc_x^{p-1}c_2 \end{matrix}$$

Breaking this up by columns into two determinants it gives

$$p\left\{(m-1)(ab)(ac)b_xc_x + (n-1)(ab)(bc)a_xc_x\right\}a_x^{m-2}b_x^{n-2}c_x^{p-2}.$$

But $$2(ab)(ac)b_xc_x = (ab)^2c_x^2 + (ac)^2b_x^2 - (bc)^2a_x^2$$

and $$2(ab)(bc)a_xc_x = -(ab)^2c_x^2 - (bc)^2a_x^2 + (ac)^2b_x^2.$$

Therefore $$2J \div (m+n-2)p$$

$$= a_x^{m-2}b_x^{n-2}c_x^{p-2}\left\{-\frac{m-n}{m+n-2}(ab)^2c_x^2 + (ac)^2b_x^2 - (bc)^2a_x^2\right\}$$

$$-\frac{m-n}{m+n-2}(f, \phi)^2\psi + (f, \psi)^2\phi - (\phi, \psi)^2 f,$$

which proves the theorem.

THEOREM II.—*The product of two binary Jacobians is reducible, for all cases excluding linear forms.*

---

[1] Cf. Gilham, *Proc. London Math. Soc.*, 2, **20** (1921–2), 326–328.

Proof.—

By use of the identity

$$(cd)a_x = (ad)c_x + (ca)d_x$$

together with the preceding identities, we have

$$\begin{aligned} 2(ab)\,(cd)a_x b_x c_x d_x &= 2(ab)\{(ad)c_x + (ca)d_x\}b_x c_x d_x \\ &= (ab)^2 c_x^2 d_x^2 + (ad)^2 b_x^2 c_x^2 - (bd)^2 a_x^2 c_x^2 \\ &\quad - (ab)^2 c_x^2 d_x^2 - (ac)^2 b_x^2 d_x^2 + (bc)^2 a_x^2 d_x^2 \end{aligned}$$

giving four terms after cancelling two. Now we multiply throughout by

$$\tfrac{1}{2} n_1 n_2 n_3 n_4 a_x^{n_1-2} b_x^{n_2-2} c_x^{n_3-2} d_x^{n_4-2}, \quad (n_i > 1),$$

and interpret the result with regard to four ground forms

$$f_1 = a_x^{n_1}, \quad f_2 = b_x^{n_2}, \quad f_3 = c_x^{n_3}, \quad f_4 = d_x^{n_4}.$$

Hence the Jacobian of f_1 and f_2, multiplied by that of f_3 and f_4 is equal to

$$\frac{n_1 n_2 n_3 n_4}{2}\{H_{14} f_2 f_3 + H_{23} f_1 f_4 - H_{13} f_2 f_4 - H_{24} f_1 f_3\}$$

where $H_{14} = (f_1, f_4)^2$, &c.; and this effects the reduction.

In these two theorems we have expressed a more complicated covariant in terms of covariants of lower degree, thereby gaining results which have wide applications. The most useful case of the latter result is when $f_1 = f_3 = f = a_x^m$, and $f_2 = f_4 = H = (ab)^2 a_x^{m-2} b_x^{m-2}$, the Hessian of f. In this case the two Jacobians in question become

$$(f_1, f_2) = (f_3, f_4) = (f, H)$$

the important covariant of degree three, usually denoted by the letter t. The theorem then tells us that the square of this covariant is reducible. On this result, or syzygy as it is called, the whole theory of solving binary cubics and quartics depends, and also the remarkable theory of finite groups of rotations whereby the five regular Platonic solids may be brought into coincidence with themselves, a subject very beautifully discussed by F. Klein.[1]

---

[1] *Lectures on the Icosahedron*, translated by G. G. Morrice (London, 1913).

9. Remarks on the Proof of the Second Fundamental Theorem.

The existing proofs for the general case of this theorem are long and difficult, but can be much simplified with the help of a recent discovery[1] by R. Weitzenböck concerning all possible relations between the compound co-ordinates $\pi_r = (xy \ldots z)_{i_1 i_2 \ldots i_r}$ where $2 \leqslant r \leqslant n-2$. For a given value of $r \leqslant \frac{1}{2}n$, every polynomial relation $R(\pi_r)$ which vanishes identically when π_r is resolved into elements $x_{i_1}, \ldots, z_{i_r}$, can be expressed as a finite series $\Sigma A_\nu \Pi_\nu$, where each Π_ν denotes a quadratic relation in π_r. Now all such p-relations as they are called fall under the type Π_2 already quoted, a well-known instance of which is

$$p_{12}p_{34} + p_{23}p_{14} + p_{31}p_{24} = 0$$

between line co-ordinates $(xy)_{ij} = p_{ij}$, when $n = 4$.

The proof is very like that of the First Fundamental Theorem, requiring also Bazin's theorem (Ex. **9**, p. 56) and the conception of implicit convolution (herausgegriffenen Reihen) for its completion.

These p-relations have hitherto been known to be sufficient to express all relations between each set of compound co-ordinates, because they furnish a particular case of the Second Fundamental Theorem. But this direct proof, recently found, provides a more direct approach to this theorem.

Another useful method is developed by B. L. van der Waerden, *Math. Annalen*, **95**, (1926), 706–736.

---

[1] Weitzenböck, *Math. Annalen*, **97**, (1927), 788-795; **99** (1928), 493–496.

CHAPTER XIV

SEMINVARIANTS. ALGEBRAICALLY COMPLETE SYSTEMS

1. Seminvariants and Leading Term of a Concomitant.

Symbolic methods lead quickly to the important result that a binary covariant is completely specified if its leading term is known, by which is meant the term with the highest power of x_1. By the fundamental theorem the symbolic product

$$C = (ab)^p (ac)^q \dots a_x^{\,r}\, b_x^{\,s} \dots$$

is a covariant, provided it contains the requisite number of symbols a, b, &c., imposed by the ground forms. If this is of order ϖ in x_1, x_2 we may adopt a new symbol and write it as α_x^{ϖ}. Thus

$$\alpha_x^{\varpi} = (ab)^p (ac)^q \dots a_x^{\,r} b_x^{\,s} \dots$$

is identically true for all values of x_1, x_2. Let $w = p + q + \dots$, the total index of the bracket factors.

Now consider the polynomial S defined by taking $x_1 = 1$, $x_2 = 0$, in C:

$$S = (ab)^p (ac)^q \dots a_1^{\,r} b_1^{\,s} \dots,$$

since in this case $a_x = a_1$, $b_x = b_1$, &c. Then S is a polynomial in the coefficients, and in fact is the leading term of the covariant C. When $x \rightarrow x'$, let $S \rightarrow S'$. Then (p. 184 (8)) since $(a'b') = (\xi\eta)(ab)$, $a_1' = a_\xi$, we have

$$\begin{aligned} S' &= (a'b')^p (a'c')^q \dots a_1'^{\,r}\, b_1'^{\,s} \dots \\ &= (\xi\eta)^w (ab)^p (ac)^q \dots a_\xi^{\,r}\, b_\xi^{\,s} \\ &= (\xi\eta)^w \alpha_\xi^{\omega}. \end{aligned}$$

If we divide S' by $(\xi\eta)^w$ we obtain α_ξ^{ϖ}, the original covariant with ξ_1, ξ_2 replacing x_1, x_2. So *from the leading coefficient of a covariant we deduce the whole covariant* by making the linear transformation on the coefficients and dividing by a suitable power, w, of the modulus. This process is non-symbolical.

For example, $(\alpha\beta)^2\alpha_x\beta_x$ is a covariant of the cubic

$$A_0x_1{}^3+3A_1x_1{}^2x_2+3A_2x_1x_2{}^2+A_3x_2{}^3=\alpha_x{}^3=\beta_x{}^3.$$

Its leading term is $(\alpha\beta)^2\alpha_1\beta_1x_1{}^2$

$$=\alpha_1\beta_1(\alpha_1{}^2\beta_2{}^2-2\alpha_1\alpha_2\beta_1\beta_2+\alpha_2{}^2\beta_1{}^2)x_1{}^2=2(A_0A_2-A_1{}^2)x_1{}^2.$$

The leading coefficient is called a *seminvariant*. Symbolically we may define a seminvariant to be a polynomial in the types (ab), a_1 where the suffix is always unity. Non-symbolically the seminvariant is usually defined to be an invariant of the *restricted* linear transformation

$$\begin{aligned} x_1 &= x_1'+\eta_1x_2', \\ x_2 &= \qquad\ \eta_2x_2', \end{aligned}$$

where it will be noted that the coefficients ξ_1, ξ_2 are 1, 0. Manifestly the symbol a_1 is now invariantive because

$$a_1'=a_\xi=a_1,$$

whereas the symbol a_2 is not. It will further be noted that this restricted transformation belongs to what has been called an *affine* group (§7, p. 162).

2. Seminvariants as Solutions of Partial Differential Equations.

The seminvariant provides a useful means of studying the concomitants of ground forms, without recourse to the symbolic methods. The leading idea which governs this alternative theory lies in the solution of a differential equation

$$a_0\frac{\partial S}{\partial a_1}+2a_1\frac{\partial S}{\partial a_2}+3a_2\frac{\partial S}{\partial a_3}+\ldots+pa_{p-1}\frac{\partial S}{\partial a_p}=0, \qquad (1)$$

where S is regarded as a polynomial in $p+1$ independent variables $a_0, a_1, \ldots, a_p$. Sylvester and Cayley first studied this equation, basing their results on the fact that if these independent variables are coefficients of a binary p-ic $(a_0, a_1, \ldots, a_p \between x_1, x_2)^p$ then every polynomial solution of the differential equation is a seminvariant of the p-ic. The seminvariant S has, in the words of Sylvester, an *annihilator* Ω, namely

$$\Omega\equiv a_0\frac{\partial}{\partial a_1}+2a_1\frac{\partial}{\partial a_2}+\ldots+pa_{p-1}\frac{\partial}{\partial a_p}. \qquad (2)$$

Further, every polynomial I annihilated both by Ω and the corresponding operator O, given by

$$O \equiv a_p \frac{\partial}{\partial a_{p-1}} + 2a_{p-1} \frac{\partial}{\partial a_{p-2}} + \ldots + pa_1 \frac{\partial}{\partial a_0} \quad . \quad (3)$$

is an invariant of the p-ic.

Granted the first result, the second follows in various ways. Thus it will be seen that O is derived from Ω by reversing the terms of the p-ic. Consequently if $\Omega S = 0$ implies S is a leading term of a covariant, then $OT = 0$ implies T is the final term of a covariant. Symbolically S being composed of types (ab), a_1 then T would involve (ab) and a_2. Accordingly they can only be the same expression if a_1, a_2 are absent and type (ab) is left, giving an invariant. Or, again, the leading term contains no x_2, so the final term contains no x_1, hence the covariant is free from both variables, and is therefore an invariant.

We therefore consider the annihilator Ω. To this end let the non-homogeneous form be taken

$$f(x) \equiv U_p = a_0 x^p + pa_1 x^{p-1} + \ldots + a_p, \quad . \quad (4)$$

so that

$$\frac{df}{dx} = pU_{p-1} = p\,(a_0, a_1, \ldots, a_{p-1} \between x, 1)^{p-1},$$

$$\frac{d^2f}{dx^2} = p(p-1)\,U_{p-2} = p(p-1)\,(a_0, a_1, \ldots, a_{p-2} \between x, 1)^{p-2},$$

and so on. Then if $x = y + h$,

$$f(y+h) = f(h) + yf'(h) + \ldots + \frac{y^p}{p!} f^{(p)}(h)$$

$$= a_0 y^p + pU_1(h) y^{p-1} + \binom{p}{2} U_2(h) y^{p-2} + \ldots + U_p(h)$$

where

$$U_1(h) = a_0 h + a_1, \quad U_2(h) = a_0 h^2 + 2a_1 h + a_2, \quad \&c. \quad . \quad (5)$$

Now if $\alpha, \beta, \ldots, \omega$ are the roots of $f(x) = 0$, then any polynomial $F(a_0, \ldots, a_p)$ is a symmetric polynomial function $\phi(\alpha, \beta, \ldots, \omega)$ of these roots. Also since $y = x - h$, the roots of the corresponding equation for y are $\alpha - h, \beta - h, \ldots, \omega - h$. So we are led to the identities

$$\left.\begin{aligned} \phi(\alpha, \beta, \ldots, \omega) &= F(a_0, a_1, \ldots, a_p) \\ \phi(\alpha - h, \beta - h, \ldots, \omega - h) &= F(a_0, U_1(h), \ldots, U_p(h)) \\ &= F(a_0, a_0 h + a_1, a_0 h^2 + 2a_1 h + a_2, \ldots). \end{aligned}\right\} \quad (6)$$

Expanding both sides as ascending power series in h and retaining only the first power of h, we have by Taylor's theorem for p variables,

$$\phi(\alpha, \beta, \ldots) - h\left(\frac{\partial}{\partial\alpha} + \frac{\partial}{\partial\beta} + \ldots + \frac{\partial}{\partial\omega}\right)\phi(\alpha, \beta, \ldots)$$

$$= F(a_0, \ldots, a_p) + h\left(a_0 \frac{\partial}{\partial a_1} + 2a_1 \frac{\partial}{\partial a_2} + \ldots + p a_{p-1} \frac{\partial}{\partial a_p}\right) F(a_0, \ldots).$$

Thus the two operators are equivalent:

$$a_0 \frac{\partial}{\partial a_1} + 2a_1 \frac{\partial}{\partial a_2} + \ldots + p a_{p-1} \frac{\partial}{\partial a_p} = -\left(\frac{\partial}{\partial\alpha} + \frac{\partial}{\partial\beta} + \ldots + \frac{\partial}{\partial\omega}\right), \quad (7)$$

the first Ω taking effect on a function F explicitly containing $a_0, \ldots, a_p$, the second on the same function expressed in terms of the roots $\alpha, \beta, \ldots, \omega$. Hence any solution of $\Omega F = 0$, when expressed in terms of the roots, is a solution of

$$\frac{\partial\phi}{\partial\alpha} + \frac{\partial\phi}{\partial\beta} + \ldots + \frac{\partial\phi}{\partial\omega} = 0. \quad (8)$$

But the general solution of this last differential equation is found from Lagrange's auxiliary equations

$$\frac{d\alpha}{1} = \frac{d\beta}{1} = \ldots = \frac{d\omega}{1} = \frac{d\phi}{0}.$$

Independent integrals of this system are $\phi =$ constant, and $p - 1$ differences of the roots. So any solution of the equation (8) is a function of the differences of the roots, say

$$\phi = S(\alpha - \beta, \alpha - \gamma, \ldots). \quad . \quad . \quad . \quad . \quad (9)$$

Again, in the alternative form $\Omega F = 0$, Lagrange's auxiliary equations are

$$\frac{da_0}{0} = \frac{da_1}{a_0} = \frac{da_2}{2a_1} = \ldots = \frac{da_p}{pa_{p-1}} = \frac{dF}{0}.$$

Independent integrals of these are $F =$ constant, and, including

a_0, p solutions involving the a's. But such solutions are given by the system

$$\begin{aligned}
S_0 &= a_0 \\
S_2 &= a_0a_2 - a_1^2 \\
S_3 &= a_0^2a_3 - 3a_0a_1a_2 + 2a_1^3 \\
S_4 &= a_0^3a_4 - 4a_0^2a_1a_3 + 6a_0a_1^2a_2 - 3a_1^4 \\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot \\
S_p &= a_0^{p-1}a_p - pa_0 \quad a_1a_{p-1} \\
&\quad + \binom{p}{2}a_0^{p-3}a_1^2a_{p-2} - \dots + (-)^{p-1}(p-1)a_1^p
\end{aligned} \tag{10}$$

For these polynomials are the numerators of the rational functions $U_0(h)$, $U_1(h)$, ..., $U_p(h)$ when $h = -a_1/a_0$, as is seen by substitution in (5). Also each $U_i(h)$ is a symmetric function of the roots of $f(y+h)=0$ as an equation in y, and further, in this particular case each root for y is a function of the differences of the roots $\alpha, \beta, \dots, \omega$ for x, since

$$\alpha - h = \alpha + \frac{a_1}{a_0} = \alpha - \frac{\Sigma\alpha}{p} = \frac{(\alpha-\beta)+(\alpha-\gamma)+\dots+(\alpha-\omega)}{p}.$$

Hence $U_i(a_1/a_0)$ is a particular function of type S. Finally since the p functions S_i are independent (S_{i+1} including a_{i+1} not found in S_i), the general solution of the equation $\Omega F = 0$ is established; namely F is an explicit function of the p polynomials S_i,

$$F(S_0, S_2, S_3, \dots, S_p).$$

Thus a polynomial in these S_i is a polynomial in the differences of roots of $f(x)=0$.

Finally, such a function satisfies the seminvariant condition; since in the present case $x = x_1 : x_2$, so that $x \rightarrow x'$ becomes

$$x = \frac{\quad\;\eta_1}{\eta_2} \qquad \text{i.e. } \eta_2 x = x' + \eta_1,$$

whence $\eta_2\alpha = \alpha' + \eta_1$ connecting any root α with α' the corresponding root of the p-ic in x'. Thus $\eta_2(\alpha-\beta) = \alpha' - \beta'$.

Accordingly if ϕ is a homogeneous function of degree w in the differences of the roots $\alpha - \beta$, the corresponding transformed function of $\alpha' - \beta'$ is ϕ multiplied by η_2^w, depending only on the coefficients of the transformation. Thus ϕ is a seminvariant and so also is its alternative form $F(S_0, \ldots, S_p)$.

It may be noticed that the degree and weight of S_i are equal to i. If S_i is the leading coefficient in a covariant of order ϖ, then (§4, p. 186) the valency condition $2w + \varpi = pq$ gives

$$2i + \varpi = ip \qquad \text{or} \qquad \varpi = i(p-2).$$

EXAMPLES

1. For a linear form, S_0 alone exists since $a_i = 0$ $(i > 1)$. For a quadratic S_0 is leader of the ground form and S_2 is the invariant discriminant.

2. Verify that $\Omega S_2 = OS_2 = 0$ for the quadratic.

3. Verify that $OS_i \neq 0$, $p > i > 2$.

4. Prove that S_2 is the leader of the Hessian covariant of a p-ic, namely,

$$\begin{vmatrix} \dfrac{\partial^2 f}{\partial x_1^2} & \dfrac{\partial^2 f}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2} & \dfrac{\partial^2 f}{\partial x_2^2} \end{vmatrix}$$

5. Show that S_3 is the leader of the Jacobian of a p-ic f and its Hessian H. We write this covariant (f, H).

6. Express S_2, S_3 in symbolic form for $f = a_x^p = b_x^p = c_x^p$.

$$S_2 = \tfrac{1}{2}(ab)^2 a_1^{p-2} b_1^{p-2}, \quad S_3 = (ab)^2 (ca) a_1^{p-3} b_1^{p-2} c_1^{p-1}.$$

7. What covariants are symbolized by

$$(ab)^2 a_x^{p-2} b_x^{p-2}, \quad (ab)^2 (ac) a_x^{p-3} b_x^{p-2} c_x^{p-1}\,?$$

8. Write down the most general polynomial of degree 3 and weight 6 in the coefficients a_0, a_1, a_2, a_3, a_4 of a binary quartic. Show that if Ω or if O annihilates it, then it is the J invariant.

Ans. $\lambda a_0 a_2 a_4 + \mu a_0 a_3^2 + \nu a_1^2 a_4 + \rho a_1 a_2 a_3 + \sigma a_2^3$, $\begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix}$

9. A homogeneous, isobaric polynomial G in $a_0, a_1, \ldots, a_p$ is a gradient. Prove for any gradient of degree q and weight w

$$(\Omega O - O\Omega)G = (pq - 2w)G.$$

10. What further condition is necessary for G to be an invariant? $pq = 2w$.

3. Algebraically Complete Systems. Syzygies.

With the help of the theory of linear partial differential equations we have found a set of p seminvariants in terms of which the most general seminvariant of a binary p-ic can be expressed. Such a set is called *algebraically complete* because it can be proved that the relations connecting other seminvariants with these p forms S_i are expressible in a rational polynomial form. They are called *syzygies.*

If S', S'' ... are seminvariants connected by a syzygy $\phi = 0$, we suppose ϕ to be isobaric in the original coefficients, otherwise it breaks up into isobaric terms each of which is itself a syzygy, exactly as was the case in §2, p. 171.

Since a covariant is determined by its leading term, it follows that any such syzygy between seminvariants leads to an exact replica between covariants (and invariants). The symbolic proof is instantaneous. For if the syzygy is written

$$\Sigma S'^{\mu_1} S''^{\mu_2} \ldots = 0,$$

and w_p is the weight of $S^{(p)}$, then the sum

$$\mu_1 w_1 + \mu_2 w_2 + \ldots$$

must be the same for every term. Hence if we put the syzygy into symbols and then change α_1 into α_x and β_1 into β_x, &c., we have

$$x_1^{\mu_1 w_1 + \mu_2 w_2 + \ldots} \Sigma C'^{\mu_1} C''^{\mu_2} \ldots = 0,$$

or

$$\Sigma C'^{\mu_1} C''^{\mu_2} \ldots = 0,$$

i.e. the covariants C', C'' ... are connected by the same relation as the seminvariants.

It follows that the theory of algebraically complete systems of concomitants can at once be derived from such a theory for seminvariants.

This theory of algebraically complete systems and of binary annihilators Ω, O, for a single ground form, can be extended.

For example,[1] if several ground forms are concerned, each has its own operators Ω and O. It can be proved that a simultaneous invariant is annihilated by the sums of these operators, namely

$$\Sigma \Omega I = 0, \quad \Sigma O I = 0.$$

A special case of this is the covariant C of one ground form, which may be regarded as an invariant of the system of this ground form with the linear form whose coefficients are $-x_2$, x_1.

---

[1] Elliott, *Algebra of Quantics* (Oxford, 1913), 120–124.

The Ω of this latter is simply $-x_2\frac{\partial}{\partial x_1}$ and its O is $-x_1\frac{\partial}{\partial x_2}$. Accordingly

$$\Sigma\Omega \equiv \Omega - x_2\frac{\partial}{\partial x_1},\quad \Sigma O \equiv O - x_1\frac{\partial}{\partial x_2}.$$

Thus a covariant is a polynomial, satisfying the valency conditions, which is annihilated by these two operators.

EXAMPLES

1. Test the Hessian $\begin{matrix} a_0x_1+a_1x_2, & a_1x_1+a_2x_2 \\ a_1x_1+a_2x_2, & a_2x_1+a_3x_2 \end{matrix}$ of the cubic $(a_0, a_1, a_2, a_3 \between x_1, x_2)^3$ with the operators $\Omega - x_2\frac{\partial}{\partial x_1}$, $O - x_1\frac{\partial}{\partial x_2}$.

2. Apply the corresponding test to the simultaneous invariant $a_0b_2 - 2a_1b_1 + a_2b_0$ of two binary quadratics.

Again, there are corresponding operators for ternary and higher categories, and analogous results giving finite algebraically complete systems, some of which have considerable importance in other branches of mathematics. Recently Forsyth[1] has given results which deal with quadratic ground forms involving one or more homogeneous sets of variables x_1, x_2, x_3, x_4. The solutions so found are the functions needed in formulating the physical invariants of the Relativity Theory.

Example.—

An algebraically complete system is that of five concomitants f_1, f_2, D_{11}, D_{12}, D_{22} of two quadratics (§5, p. 219). For the sixth J can be expressed algebraically but irrationally in terms of them as

$$J = (D_{12}f_1f_2 - \tfrac{1}{2}D_{11}f_2^2 - \tfrac{1}{2}D_{22}f_1^2)^{\frac{1}{2}},$$

owing to the syzygy

$$2J^2 = 2D_{12}f_1f_2 - D_{11}f_2^2 - D_{22}f_1^2$$

which connects them. Or again, this relation might be used to give one invariant, D_{12} or D_1 or D_2, *rationally* but not integrally in terms of the other five forms.

4. **Irreducibility. Gordan's Theorem.**

The question now arises, are we to break the integrity of our work by introducing an awkward irrationality by solving this equation for J? The instinct of all the great algebraists of last century has been to say, No. Far more is gained by retaining the set of six polynomial functions than by rejecting one of them at the expense of symmetry. We are here in touch with a big question, one which is not often explained and reasoned out in an English elementary textbook. Perhaps our national love of independence is at work, and we unconsciously admire p things essentially distinct, rather than $p+q$ things dependent on q binding relations! Anyhow we have to thank Cayley and Sylvester for their original handling of the dilemma, with their insistence on maintaining the polynomial character of these functions at all costs. So we are led, from the algebraically complete system of concomitants, to the broader conception of the *irreducible system*, and with it, Gordan's theorem.

---

[1] *Proc. Royal Soc. Edinburgh,* **42** (1921–2), 147–212.

Because every polynomial concomitant of two binary quadratics can be expressed rationally and integrally in terms of six, but not of five or less of them, these six are called the irreducible system of two quadratics. Analogous systems hold for a single binary cubic, quartic, and n-ic.

For a time after Cayley first broached the subject in 1856 the conviction began to gain ground that for values of n greater than four the system was infinite. Then in 1868 Gordan sprang his great surprise on the algebraic world by proving that the irreducible system of a binary n-ic is *finite*; and this in short is his great theorem. He perfected the proof in three stages, extending it from the original case of one to any number of binary ground forms.[1]

In 1890 Hilbert[2] gave an alternative proof applicable to forms of all categories. This proof, which we shall soon consider in detail, consists of two parts, first establishing a remarkable, not to say startling, lemma, sometimes called the Basis Theorem, of very great generality; and secondly leading by use of the Cayley operator to the desired result concerning invariants. Hilbert's lemma is an Existence Theorem: it establishes the existence of a certain finite system of functions, but throws no light on how to find them. Gordan's proof, on the other hand, actually provides its own solution. Both methods were very great achievements.

---

[1] Cayley, *Second Memoir* (1856). *Collected Papers*, Vol. II, 250–275. Gordan, *Crelle*, **69** (1868), 323–354. On p. 343 the author introduces the term *complete system* (*volles System*). Grace and Young, *Algebra of Invariants*, pp. 101–127 contain the proof for binary forms, substantially in the form of Gordan's third proof. For a general survey of the whole problem, see Meyer's *Berichte*, pp. 134–150.

[2] *Math. Annalen*, **36** (1890).

CHAPTER XV

THE GORDAN-HILBERT FINITENESS THEOREM

1. Hilbert's Basis Theorem.

Let

$$X_1,\ X_2,\ \ldots,\ X_n \qquad \ldots\ldots \qquad (1)$$

be n variables. Further, let there be a given law or specified set of conditions whereby *forms* F in these n variables are constructed. Let this law be such that it leads to an infinite number of such forms, each involving the variables and not being merely a constant. We write

$$S_\infty = F_1,\ F_2,\ \ldots,\ F_m,\ \ldots \qquad \ldots \qquad (2)$$

to denote the totality S_∞ of these forms F_i.

Suppose further that

$$A_1,\ A_2,\ \ldots,\ A_m \qquad \ldots\ldots \qquad (3)$$

denote forms in X, which are not necessarily contained in the system S_∞, but which can if necessary be constants. Then Hilbert's Basis Theorem runs as follows:

From the infinite set of forms S_∞ *a finite set of forms* F_1, $\mathrm{F}_2, \ldots,$ F_m *can be selected, such that every form* F *of the system* S_∞ *can be expressed as*

$$F \equiv A_1F_1 + A_2F_2 + \ldots + A_mF_m. \qquad \ldots \qquad (4)$$

These forms $F_1, F_2, \ldots, F_m$ are then said to be the *basis* of the system S_∞.

To illustrate this we can take the extreme case when the law is absolutely general. In this case $F_1, F_2, \ldots, F_m$ are the n variables $X_1, X_2, \ldots, X_n$ themselves. For each of these is a linear form, and thus falls within the set S_∞; while any other form F can manifestly be expressed as

$$A_1X_1 + A_2X_2 + \ldots + A_nX_n,$$

where each A_i is a suitable form of one degree lower than that of F.

This basis would fail if the law specifically excluded linear forms, because none of X_i would then fall within S_∞.

Proof.—

We prove the theorem by induction. For if $n = 1$, there is only one variable X_1 and the most general homogeneous polynomial, or form, F_i is now simply $c_i X_i^{k_i}$. Then the supposed law will produce various positive integral indices k_i, of which k_1, say, is the least or one of the least. Every form is then expressible as in (4) with $m = 1$, since F_1 will now be a factor of every possible form F_i. So the theorem is true if $n = 1$.

Next let us assume it true for $n - 1$ variables $X_1, X_2, \ldots, X_{n-1}$. Let F_1, an arbitrary member of the system S_∞, be of degree r in the n variables $X_1, X_2, \ldots, X_n$. It is then quite possible that F_1 contains no term involving X_n^r, for the supposed law might expressly exclude this. We can, however, guarantee that such a term actually occurs, a fact of importance later in the proof, by the following device.

Through the linear transformation $X \rightarrow Y$, or more expressly,

$$X_i = \sum_j e_{ij} Y_j, \qquad \Delta = | e_{ij} | \neq 0, \quad . \quad . \quad . \quad (5)$$

we change the form F_1 into a form G_1, of degree r in the variables Y; say

$$G_1 = c Y_n^r + c_1 Y_n^{r-1} + \ldots + c_r \qquad (c \neq 0). \quad . \quad (6)$$

The constant coefficient c of Y_n^r, which alone we need to examine, is given by substituting the values $X_i = e_{in}$ $(i = 1, 2, \ldots, n)$ in F_1 (§2, pp. 183–4); in fact

$$c = (F_1)_{X_i = e_{in}}:$$

and since F_1 is not identically zero, we may choose suitable values of these coefficients e_{in} so as to give $c \neq 0$.

Next if we transform members F_i of S_∞, by (5), into forms $G_1, G_2 \ldots$ in the variables Y, we can prove the theorem for the forms F by showing that it holds for G. For if

$$G = B_1 G_1 + B_2 G_2 + \ldots + B_m G_m \quad . \quad . \quad (7)$$

holds for every form G in terms of a basis $G_1, \ldots, G_m$, we merely have to reverse the transformation $X \rightarrow Y$ to obtain the desired result (4) directly from (7). We therefore proceed to prove the result (7), taking G_1, as given by (6).

To this end let each G_s $(s = 2, 3, \ldots)$ be divided by G_1 regarded as a function of Y_n; giving a quotient B_s. Thus

$$G_s = B_s G_1 + \Phi_s Y_n{}^{r-1} + \Psi_s Y_n{}^{r-2} + \ldots + \Omega_s. \qquad (8)$$

Here B_s is identically zero if G_s is of lower degree in Y_n than G_1: also some or all of the polynomials $\Phi_s, \Psi_s, \ldots, \Omega_s$ may vanish; and further, they all may contain $n-1$ variables $Y_1, Y_2, \ldots, Y_{n-1}$ but not Y_n. To such polynomials, formed as they are by a definite law, the Basis Theorem applies. In particular for the system

$$\Phi_2, \Phi_3, \ldots, \Phi_s, \ldots$$

we have

$$\Phi_s = A_{s_1}\Phi_{s_1} + A_{s_2}\Phi_{s_2} + \ldots + A_{s_\nu}\Phi_{s_\nu}, \quad . \quad . \quad (9)$$

where the suffixes $s_1, s_2, \ldots, s_\nu$ are fixed integers.

If we write (8) as

$$\Phi_s Y_n{}^{r-1} = G_s - B_s G_1 - \Psi_s Y_n{}^{r-2} - \ldots - \Omega_s$$

and let s take the ν special values $s_1, s_2, \ldots, s_\nu$, we obtain the system of equations

$$\left.\begin{array}{l} \Phi_{s_1} Y_n{}^{r-1} = G_{s_1} - B_{s_1} G_1 - \Psi_{s_1} Y_n{}^{r-2} - \ldots - \Omega_{s_1} \\ \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ \Phi_{s_\nu} Y_n{}^{r-1} = G_{s_\nu} - B_{s_\nu} G_1 - \Psi_{s_\nu} Y_n{}^{r-2} - \ldots - \Omega_{s_\nu} \end{array}\right\}. \quad (10)$$

Multiplying these respectively by $A_{s_1}, A_{s_2}, \ldots, A_{s_\nu}$ and adding, we obtain, by (9),

$$\begin{aligned} \Phi_s Y_n{}^{r-1} = {} & (A_{s_1}G_{s_1} + \ldots + A_{s_\nu}G_{s_\nu}) \\ & + CG_1 + \Psi_s' Y_n{}^{r-2} + \ldots + \Omega_s', \quad . \quad . \quad (11) \end{aligned}$$

where the significance of $C, \Psi_s', \ldots, \Omega_s'$ is easily seen. Whence by substituting this value of $\Phi_s Y_n{}^{r-1}$ in (8), we have, say,

$$\begin{aligned} G_s &= B_s' G_1 + A_{s_1}G_{s_1} + \ldots + A_{s_\nu}G_{s_\nu} + \Psi_s'' Y_n{}^{r-2} + \ldots + \Omega_s'' \\ &= \Sigma B_i G_i + \Psi_s'' Y_n{}^{r-2} + \ldots + \Omega_s'' \\ &= \sum_i B_i G_i + U_{r-2}, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (12) \end{aligned}$$

where i takes $\nu + 1$ values, $1, s_1, s_2, \ldots, s_\nu$. In the same way (8) may be written $G_s = B_s G_1 + U_{r-1}$, where U_{r-1}, U_{r-2} are polynomials of degree in Y_n at most indicated by their suffixes. Proceeding again in this way as from (8) to (12), we deduce an expression for G_s involving U_{r-3}, of degree at most $r-3$. This procedure, in at most $(r-1)$ steps, gives U_0, a polynomial independent of Y_n, and therefore one for which the basis theorem is true. So finally we express G_s as

$$G_s = B_1 G_1 + B_2 G_2 + \ldots + B_m G_m,$$

where m is finite. This proves the theorem.

The remarkable point about this theorem is its breadth: for provided the law of formation of the functions F_i is definite it may be as intricate a law as we please. For instance, the variables $X_1, X_2, \ldots, X_n$ may be replaced by any finite set of variables, such as the cogredient and contragredient and intermediate compound sets x, u, p_{ij}, π^{ij}, &c., already adopted.

Further, the proof is applicable not merely to the field of ordinary complex numbers, but to any more restricted ring (p. 12) of number where the law of division, with remainder, as required in (8), together with laws of addition and multiplication for constructing polynomials, still hold.[1]

2. Proof of Gordan's Theorem.

We can now give a formal proof of the Gordan-Hilbert Finiteness Theorem, which runs as follows:

For any finite given set of ground forms, every rational integral concomitant of a general linear transformation can be expressed rationally and integrally in terms of a finite number of concomitants $C_1, C_2, \ldots, C_m$. *These* m *concomitants are said to form the* **complete system** *for the given ground forms.*

Proof.—

By adjoining certain linear forms to the given ground forms every concomitant of the original system is determined by the invariants of this extended system (§8, p. 207). We therefore need consider invariants only, these being functions of the coefficients of the ground forms.

Since homogeneous invariants represent all polynomial invariants (§3, p. 171), while the multiplication of a given invariant throughout by a non-zero constant does not effectively change it, we may consider all invariants as *forms* in the coefficients of the ground forms, derived therefrom by definite laws. Thus the Hilbert Basis Theorem applies to any invariant I; namely

---

[1] The above is substantially the original proof given by Hilbert. An even neater proof was given later by Gordan (*Göttinger Nachrichten* (1899), 240–242). Cf. Grace and Young, *Invariants*, pp. 178–182.

$$I = A_1 I_1 + A_2 I_2 + \dots + A_m I_m, \qquad . \quad . \quad (13)$$

where $I_1, I_2, \dots . I_m$ is a fixed set of invariants, and each A_i is a polynomial in the coefficients of the ground forms, but not necessarily an invariant.

Now let the linear transformation of variables change A_i to A_i', I_k to $(\xi\eta\dots\omega)^{w_k} I_k$, and I to $(\xi\eta\dots\omega)^w I$. Then the kth term on the right of (13) is transformed to a polynomial in $\xi_1, \dots, \omega_n$ of weight w; in order to agree with the left. Further it breaks up into, say, r single terms (if A_k has r terms), each of which has the factor I_k independent of $\xi_1, \dots, \omega_n$. All other factors can be symbolized by inner and outer products $a_\xi, \dots$, $(\alpha\eta\zeta\dots\omega), \dots$, with a possible common denominator $(\xi\eta\dots\omega)^\rho$, exactly as in §7, p. 205. We multiply throughout by this denominator, so that (13) is replaced by an identity

$$(\xi\eta\dots\omega)^{w+\rho} I = \phi_1 I_1 + \phi_2 I_2 + \dots + \phi_m I_m, \qquad w + \rho > 0,$$

where each ϕ_i is a polynomial in these inner and outer products.

Operating with $\Omega^{w+\rho}$ on both sides as in §7, p. 206, we obtain

$$\lambda I = I_{\mu_1} I_1 + I_{\mu_2} I_2 + \dots + I_{\mu_m} I_m, \qquad (\lambda \neq 0), \qquad (14)$$

where the kth term $I_{\mu_k} I_k$ is due to the kth term of the original series, I_{μ_k} being either an invariant or zero.

Since I_k has degree greater than zero, I_{μ_k} has degree less than that of I. Thus every invariant is expressible polynomially in terms of $I_1, \dots, I_m$, together with invariants I_{μ_k} of lower degree. Treating each I_{μ_k} by the same process as for I itself, we lower the degree of these additional invariants at each stage. Since the degree of I is finite, a finite number of such processes ultimately furnishes mere constants, apart from the system $I_1, \dots, I_m$. So we have expressed I explicitly as a polynomial in $I_1, \dots, I_m$; and this proves the theorem. A proof without the help of symbolic methods can also be given.[1]

3. Limit to the Number of Syzygies.

The distinction between an algebraically complete and an irreducibly complete system for given ground forms should now be clear. Manifestly the latter is richer and more elaborate than the former, as, for instance, in the case of two binary quadratics (10), p. 220, where a syzygy connects the six irreducibles so as to make five algebraically independent. Evidently, too, there will be such syzygies in general between members of the irreducible system.

---

[1] Cf. Weitzenböck, *Invariantentheorie* (1923), 145–148.

It is an interesting example of the Basis Theorem to prove that for a given complete system of m forms

$$I_1, I_2, \ldots, I_m,$$

the number of polynomial syzygies is limited.

For let $G(I) = 0$ be called a *syzygy of the first kind* if $G(I)$ is a polynomial in these I's which does not vanish identically, and which only involves coefficients of the ground forms implicitly by means of these I's, and which vanishes when each I is expanded as a polynomial in these coefficients.

Then $G(I)$ can if necessary be made homogeneous by adjoining another I_0 which later can be made equal to unity. To such functions, built by definite laws, the Basis Theorem applies. Hence every G is expressible as

$$G(I) = A_1 G_1(I) + \ldots + A_\nu G_\nu(I)$$

in terms of a finite number of such functions.

4. Multiple Fields.

Hilbert has pointed out the applicability of his methods to the general theory of forms when the variables fall into quite distinct fields governed by independent linear transformations. The following example[1] sufficiently illustrates the method, which may readily be generalized.

Let
$$F = A_X{}^p a_x{}^q \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (15)$$

be a form of orders (p, q) in two independent sets of variables X and x:

$$X_1, X_2, \ldots, X_m; \; x_1, x_2, \ldots, x_n, \quad . \quad . \quad (16)$$

where X belongs to a field of order m, and x to one of order n. Here m and n may be equal or not as we prefer.

The symbolic coefficient of a typical term is

$$A_i A_j \ldots a_r a_s \ldots \quad . \quad . \quad . \quad . \quad . \quad . \quad (17)$$

---

[1] Cf. Weitzenböck, *Invariantentheorie*, p. 169.

involving p factors with a capital A and q with a small a. The actual coefficient could be represented by

$$A_{ij\ldots,} \tag{18}$$

without recourse to upper suffixes, since the variables X, x are not assumed to be contragredient.

For such forms we require two independent linear transformations $X \rightarrow X'$ and $x \rightarrow x'$, associated with which there will be polynomial invariants composed of coefficients $A_{ij\ldots,rs\ldots}$.

If $$X_i = \sum_j E_{ij} X_j', \qquad x_i = \sum_j e_{ij} x_j' \tag{19}$$

are the linear transformations, with moduli

$$\Delta = |E_{ij}| \neq 0, \qquad \delta = |e_{ij}| \neq 0, \tag{20}$$

then as in §2, p. 170, an invariant satisfies the condition of transformation

$$I' = \Delta^W \delta^w I, \quad . \quad . \quad . \quad . \quad . \quad . \quad . \tag{21}$$

where W, w, positive integers, are called the weights of I in the coefficients of the independent sets of variables.

For such polynomials in coefficients (18) the two Cayley operators

$$\Omega_E = \genfrac{}{}{0pt}{}{}{\partial E_{ij}} \qquad \Omega_e = \Big| \genfrac{}{}{0pt}{}{\partial}{\partial e_{ij}} \tag{22}$$

have place. A symbolic polynomial P in the symbols A, a, which contains after transformation a weight W of the E_{ij} and w of the e_{ij} is such that

$$\Omega_E{}^W \Omega_e{}^w P = 0 \quad \text{or} = I \quad . \quad . \quad . \quad . \tag{23}$$

an invariant. Hence as in the proof of the First Fundamental Theorem an invariant is symbolized by aggregates of two sorts of bracket factors

$$(AB \ldots M), \qquad (ab \ldots n) \quad . \quad . \quad . \tag{24}$$

which are symbolic determinants of orders m and n respectively. A covariant will also have factors A_X, a_x.

Further, the Hilbert Basis Lemma applies; and, for a polynomial invariant we have

$$I = A_1 I_1 + A_2 I_2 + \ldots + A_\mu I_\mu$$

in terms of a properly chosen finite set of invariants $I_1, I_2, \ldots, I_\mu$. Finally the operator $\Omega_E{}^W \Omega_e{}^w$, acting upon this, leads as before to the expression of all invariants in terms of a complete system consisting of these μ invariants $I_1, I_2, \ldots, I_\mu$.

5. Combinants.

In a multiple field let F_λ be a form *linear* in one set of variables $\lambda_1, \lambda_2, \ldots, \lambda_m$. Then to take the case of two sets of variables λ, x, the typical form has orders $(1, p)$, namely

$$F_\lambda = r_\lambda a_x{}^p = \Sigma\, A_{i_1 i_2 \ldots i_{p+1}} \lambda_{i_1} x_{i_2} \ldots x_{i_{p+1}}.$$

It is simplest to express this explicitly as a sum of m terms

$$\lambda_1 f_1 + \lambda_2 f_2 + \ldots + \lambda_m f_m,$$

where each f_i is a form of order p in the variables x. Any concomitant of the multiple form F_λ is called a *combinant* of the m forms f_i.

If the independent transformations $\lambda \rightarrow \lambda'$, $x \rightarrow x'$ are made in the particular case when $\lambda' = \lambda$, then x alone changes, and the concomitant C is therefore invariantive for the form F_λ regarded as a function of the x's alone. Hence it is an invariant of the simultaneous system $f_1, f_2, \ldots, f_m$. Every combinant is therefore a concomitant of the forms f_i, but the converse is not true. Indeed it is an important problem for a given set of p-ics f_i, to determine which of their concomitants are combinants.

Similar remarks apply to cases with more sets of variables $x, y, \ldots$.

Example.—

The Jacobian of two binary forms

$$f_1 = a_x{}^p, \quad f_2 = b_x{}^p$$

is a combinant, since

$$\begin{aligned} \frac{\partial(\lambda_1 f_1 + \lambda_2 f_2,\ \mu_1 f_1 + \mu_2 f_2)}{\partial(x_1, x_2)} &= \begin{matrix} \lambda_1 & \mu_1 \\ \lambda_2 & \mu_2 \end{matrix}\ \frac{\partial(f_1, \quad f_2)}{\partial(x_1, \quad x_2)} \\ &= p^2 (\lambda\mu)(ab) a_x{}^{p-1} b_x{}^{p-1}. \end{aligned}$$

The general properties of binary combinants were given by Gordan, *Math. Annalen*, **5** (1872), p. 95. Further references are to be found in Meyer's *Berichte*, and the same author's book, *Apolarität und rationale Curve*. Much interesting information will also be found in Chapters XI and XIV, Grace and Young, *Algebra of Invariants*.

EXAMPLES[1]

1. Extend the methods of §5, p. 173, to show that a double binary form can be symbolized as follows:

$$\Psi = \sum_{i,j} \binom{m}{i} \binom{m}{j} a_{ij} x_1^{m-i} x_2^{i} y_1^{n-j} y_2^{j} = (a_1 x_1 + a_2 x_2)^m (\alpha_1 y_1 + \alpha_2 y_2)^n$$

so that $a_{ij} = a_1^{m-i} a_2^{i} \alpha_1^{n-j} \alpha_2^{j}$.

2. If this form is written $\Psi = a_x^m \alpha_y^n = b_x^m \beta_y^n$, prove that

$$(ab)(\alpha\beta) a_x^{m-1} b_x^{m-1} \alpha_y^{n-1} \beta_y^{n-1}$$

is a covariant for independent linear transformation $x \longrightarrow x'$, $y \longrightarrow y'$.

3. If $z = \xi + i\eta$, $z' = \xi - i\eta$ are conjugate complex numbers, where ξ, η are rectangular Cartesian co-ordinates of a point, then the equation of a circle can be written

$$azz' + bz + b'z' + d = 0.$$

Determine the conditions for the equation in ξ, η to have real coefficients.
[Coefficients a, d real; b, b' conjugate complex numbers.]

4. If, further, $z = x_1 : x_2$, $z' = y_1 : y_2$, $m = n$, then the equation $\Psi = 0$, expressed in terms of ξ, η, represents a plane curve of order m, in general with multiple points of order m at the circular points at infinity.
[Terms of highest degree are $(\xi^2 + \eta^2)^m$].

5. Prove that the linear transformation $w = (pz + q)/(rz + s)$, where $ps - qr \neq 0$ is equivalent to a particular case of the two transformations $x \longrightarrow x'$, $y \longrightarrow y'$ above; and that this z transformation is equivalent to inversion successively in circles orthogonal to each other.

6. If the bilinear form $(m = n = 1)$ represents a circle, prove that it degenerates to a point circle when $(ab)(\alpha\beta) = 0$.

---

[1] References to double binary forms: Peano first gave the system for bilinear forms $(m = n = 1)$; *Battaglini*, 20 (1882). For a more direct proof cf. *Proc. Roy. Soc. Edinburgh*, **43** (1922–3), 43–50 (45). The general theory is given by Kasner, *Trans. American Math. Soc.*, **1** (1900): "The invariant theory of the inversion group"; also **4** (1903).

Peano gave the 18 concomitants of the complete system of the (2, 2) form. For their geometrical treatment cf. Turnbull, *Proc. Edinburgh Roy. Soc.*, **44** (1923–4), 23–50 where other references are given; and Vaidyanathaswamy, *Proc. London Math. Soc.*, **2, 24** (1925), 83–102, "On the rank of the double binary form".

The (2, 1) form has been treated by these authors in the works quoted, while the system for two (2, 1) forms is given by Saddler, *Proc. Edinburgh Roy. Soc.*, **45** (1924–5), 3–13; cf. also **46** (1925–6), 264–282. The same author gives the system of the (1, 1, 1) form in *Proc. Cambridge Phil. Soc.*, **22** (1923–5), 688–693: cf. also Schwartz, *Math. Zeitschrift*, **12** (1922), 18–35.

For a proof of Gordan's theorem and a general transvectant method of discussing the double binary forms, cf. *Proc. Edinburgh Math. Soc.*, **41** (1922–3), 116–127: cf. also Gordan *Math. Annalen*, **33** (1889), 387–389; Study, *Math. Annalen*, **27** (1886); Lehnen, *Dissertation Bonn* (1921). Double and multiple binary perpetuants are considered by the author in *Proc. London Math. Soc.*, **2, 27** (1928), 193–208.

7. Prove that if a, x; b, β; c, γ; d, δ are pairs of symbols not necessarily equivalent, the product of two covariants

$$(ab)\,(\alpha\beta)\,a_x^{m-1}\,b_x^{m-1}\,\alpha_y^{n-1}\,\beta_y^{n-1}\,(cd)\,(\gamma\delta)\,c_x^{m-1}\,d_x^{m-1}\,\gamma_y^{n-1}\,\delta_y^{n-1}$$

is reducible. [Cf. p. 224.

6. Further Examples of Complete Systems. The Binary Cubic.

In binary forms the complete system of a single n-ic includes invariants and covariants. For a linear form ($n = 1$) the Fundamental Theorem shows that no concomitant exists beyond the form itself. A binary quadratic $f = a_x^2 = b_x^2$ has a complete system of two forms: f and its discriminant. That of a binary cubic

$$f = a_x^3 = b_x^3 = a_0 x_1^3 + 3a_1 x_1^2 x_2 + 3a_2 x_1 x_2^2 + a_3 x_2^3 \qquad (25)$$

has been established in various ways. It consists of four forms

$$\begin{gathered} f = a_x^3, \qquad H = (ab)^2 a_x b_x, \qquad t = (ab)^2 (ca)\, b_x c_x^2, \\ \Delta = (ab)^2\,(ac)\,(bd)\,(cd)^2. \end{gathered} \qquad (26)$$

Here H is the Hessian of the cubic ground form f; t is the Jacobian of f and H; and Δ is the single invariant, the discriminant of the quadratic H.

Non-symbolically we find

$$H = 2\begin{vmatrix} a_0 x_1 + a_1 x_2 & a_1 x_1 + a_2 x_2 \\ a_1 x_1 + a_2 x_2 & a_2 x_1 + a_3 x_2 \end{vmatrix} \qquad (27)$$

$$t = (f, H) = \tfrac{1}{6}\begin{vmatrix} \dfrac{\partial f}{\partial x_1} & \dfrac{\partial H}{\partial x_1} \\ \dfrac{\partial f}{\partial x_2} & \dfrac{\partial H}{\partial x_2} \end{vmatrix} \qquad (28)$$

$$\Delta = 2\left\{4\,(a_0 a_2 - a_1^2)\,(a_1 a_3 - a_2^2) - (a_0 a_3 - a_1 a_2)^2\right\}. \quad (29)$$

The leading term in t is

$$(a_0^2 a_3 - 3a_0 a_1 a_2 + 2a_1^3)\, x_1^3,$$

which contains a seminvariant of degree and weight 3 for its coefficient.

These four forms f, H, t, Δ are irreducible but not algebraically independent, for they are connected by the syzygy

$$2t^2 + H^3 + \Delta f^2 = 0, \qquad (30)$$

which is a further example of the general fact that the square of a Jacobian is reducible. This syzygy is also deducible by eliminating x_1, x_2 from the three equations

$$\begin{aligned} f &= a_0 x_1^3 + \ldots, \\ H &= 2(a_0 a_2 - a_1^2) x_1^2 + \ldots, \\ t &= (a_0^2 a_3 - 3a_0 a_1 a_2 + 2a_1^3) x_1^3 + \ldots, \end{aligned} \qquad \ldots \quad (31)$$

which is an example showing that an invariant can often be looked upon as the result of eliminating p unknowns from $p + 1$ equations.

EXAMPLE

If M is the modulus of the transformation $x_1, x_2 \longrightarrow X, Y$, where X, Y are the linear factors of the Hessian, show that this system for the cubic can be written

$$f = X^3 + Y^3, \quad H = 2M^2XY, \quad t = M^3(X^3 - Y^3), \quad \Delta = -2M^6,$$

and verify the syzygy.

7. The Binary Quartic Form.

The complete system of a binary quartic form

$$\begin{aligned} f &= a_x^4 = b_x^4 = c_x^4 \\ &= a_0 x_1^4 + 4a_1 x_1^3 x_2 + 6a_2 x_1^2 x_2^2 + 4a_3 x_1 x_2^3 + a_4 x_2^4 \end{aligned} \qquad (32)$$

consists of five concomitants

$$\begin{aligned} f = a_x^4,\ H = (f, f)^2 = (ab)^2 a_x^2 b_x^2,\ t = (f, H) = (ab)^2(ca) a_x b_x^2 c_x^3, \\ i = (ab)^4, \quad j = (bc)^2(ca)^2(ab)^2, \end{aligned} \qquad \ldots \quad (33)$$

three being covariants and two invariants.

Non-symbolically the invariants are

$$i = 2(a_0 a_4 - 4a_1 a_3 + 3a_2^2),$$

$$\begin{aligned} j &= 6 \begin{vmatrix} a_0 & a_1 & a_2 \\ a_1 & a_2 & a_3 \\ a_2 & a_3 & a_4 \end{vmatrix} \\ &= 6(a_0 a_2 a_4 - a_0 a_3^2 - a_1^2 a_4 - a_2^3 + 2a_1 a_2 a_3). \end{aligned} \qquad (34)$$

The Hessian H is a quartic of degree two in the coefficients a_i of the ground form, while t is a sextic. These two covariants are evidently analogous to those of the previous cubic ground form, as the symbolic expression shows. Corresponding covariants naturally exist for the general n-ic. The invariants i and j appear for the first time because they involve four symbols a in their bracket factors. They lead to analogous *covariants* of a quintic

$$(ab)^4 a_x b_x, \quad (bc)^2(ca)^2(ab)^2 a_x b_x c_x,$$

and for higher forms.

Between these five concomitants of a quartic a syzygy exists

$$2t^2 = \tfrac{1}{2} iHf^2 - H^3 - \tfrac{1}{3} jf^3, \quad . \quad . \quad . \quad (35)$$

again because the square of the Jacobian t is reducible. This may be verified by applying the general theorem, or by use of a canonical form, say

$$f = X^4 + 6mX^2Y^2 + Y^4. \quad . \quad . \quad . \quad (36)$$

EXAMPLE

Assuming a linear transformation $x \longrightarrow X$ of modulus $M \neq 0$ gives $a_0 x_1^4 + \ldots + a_4 x_2^4 = f = X^4 + 6mX^2Y^2 + Y^4$, prove that

$$H = 2M^2 m \left(X^4 + \frac{1 - 3m^2}{m} X^2 Y^2 + Y^4 \right),$$
$$t = M^3(1 - 9m^2) XY(X^4 - Y^4),$$
$$i = 2M^4(1 + 3m^2), \quad j = 6M^6(m - m^3),$$

and verify the syzygy (35).

8. **References to Complete Systems.**

For canonical forms when the cubic or quartic ground form is special, the reader should consult Elliott's *Algebra of Quantics*. The corresponding symbolic forms are given by Grace and Young in *The Algebra of Invariants*, where also an account of the complete systems of the binary quintic, the sextic, two cubics, quadratic and cubic, quadratic and quartic, and also of two ternary quadratics, will be found.

The septimic and octavic were worked out by v. Gall, *Math. Annalen*, **31** (1888), p. 318 and **17** (1881), p. 31, p. 139.

All these results beyond the quartic case are very complicated. There are, for example, 23 irreducible concomitants of the binary quintic. This increase of complexity is not entirely due to the increase in number of coefficients of the ground form as its order advances, for certain concomitants actually are reducible for higher orders, even when irreducible for lower. Thus the invariant

$$\Delta = (ab)^2 (ac) (bd) (cd)^2$$

of a cubic is irreducible, whereas the corresponding covariant

$$\Delta' = (ab)^2 (ac) (bd) (cd)^2 a_x b_x c_x d_x$$

of a quartic can in fact be reduced to a linear combination of jf and iH.

There is a theory of *perpetuants* which deals with covariants of a given degree for forms of order not less than the weight of any such covariant. It may be regarded as the theory of binary forms of infinite order. It affords a notable example of the value of both symbolic and non-symbolic methods of attack. For the complete system of such forms may be said to be known. It will be seen from the examples of symbolic methods in §3, p. 216, that any such system so found is comprehensive: all irreducible forms are certainly included. But it may contain redundant forms. Now the non-symbolic methods proceed in just the converse way, and show that any system so found contains no reducible terms. When the two methods yield the same result, as in the case of the binary quintic or binary perpetuants, they therefore confirm each other.

In higher fields complete systems are known in certain cases, but apart from linear and quadratic cases the only complete *ternary* system actually computed is that of the cubic by Clebsch and Gordan (*Math. Annalen,* **6** (1875), 436). The ternary quartic has received much attention but still remains unworkable.[1] The problem of ternary perpetuants was solved by Dr. A. Young (*Proc. London Math. Soc.,* **2**, **22** (1922–3), 171–220.

---

[1] A notable instalment was worked by Fräulein E. Noether, *Crelle,* **134** (1908), 23–94.

CHAPTER XVI

CLEBSCH'S THEOREM

1. Introduction of Clebsch's Theorem.

The object of the present chapter is to develop the general invariant theory as far as the variables are concerned, and the principal result will be a theorem due to Clebsch which tells us that a completely adequate account of concomitants in the field of order n can be given by restricting the choice of ground forms to functions of at most $n-1$ sets of cogredient variables. All other sets which enter can be accounted for by polarization, or by E the absolute concomitant of the field.

For example, in the binary field ($n=2$) the bilinear form $a_x b_y$ may be written

$$\begin{aligned} a_x b_y &= \tfrac{1}{2}(a_x b_y + a_y b_x) + \tfrac{1}{2}(a_x b_y - a_y b_x) \\ &= \tfrac{1}{2}\left(y \frac{\partial}{\partial x}\right) a_x b_x + \tfrac{1}{2}(ab)\,(xy). \end{aligned} \quad . \quad . \quad (1)$$

Here the first term is a polar of $a_x b_x$ which contains only one variable, and the second is a product of an invariant (ab) and the absolute covariant (xy). Further, this invariant belongs to the ground form $a_x b_x$, without the need of the second variable y.

On the other hand, in the ternary or any higher field ($n > 2$) we obtain

$$a_x b_y = \tfrac{1}{2}\left(y \frac{\partial}{\partial x}\right) a_x b_x + \tfrac{1}{2}(ab \mid xy), \quad . \quad . \quad (2)$$

where now the second term is irreducible, and involves a new type of variable $(xy)_{ij}$ which cannot be overlooked; nor could it arise if merely one set of variables x was utilized. Also the function $(ab \mid xy)$ is not a polar, but, rather, satisfies the differential equation

$$\left(y \frac{\partial}{\partial x}\right)\phi \equiv \sum_{i=1}^{n} y_i \frac{\partial}{\partial x_i} \phi = 0 \quad . \quad . \quad . \quad . \quad (3)$$

as is immediately apparent since

$$\left(y\frac{\partial}{\partial x}\right)(ab\mid xy)=(ab\mid yy)=0. \tag{4}$$

2. Compound Polars. Standard Forms.

Suppose we have any number of point co-ordinates x, y, z, ..., ξ, ... from which the various compound co-ordinates π_2, π_3, ... are derived as in §8, p. 86. We take

$$\pi_2=\overline{xy},\quad \pi_3=\overline{xyz},\quad \ldots,\quad \pi_{n-1}=\overline{xyz\ldots s}$$

for the second, third, ..., $(n-1)$th compounds, the last representing a set of n homogeneous prime co-ordinates u.

Manifestly a form involving these sets x, π_2, ..., π_{n-1} as sole variables is a polynomial function of the $n-1$ cogredient variables x, y, z, ..., s, so that we can write it

$$F(x,\pi_2,\pi_3,\ldots,\pi_{n-1})=f(x,y,z,\ldots,s).$$

But the converse is not necessarily true, as the single example already cited shows. Yet if we introduce any number of cogredient compound variables ρ_2, σ_2, ..., ρ_{n-1}, σ_{n-1}, ... and polarize each such form F in every possible way, we obtain a wider choice of forms, which may be typified by ΔF, such that every form $f(x, y, \ldots)$ can be expressed in terms of these ΔF: and this in fact is Clebsch's theorem.

By such polarization is meant an operation on F of one of the following types:

$$\left(\xi\frac{\partial}{\partial x}\right),\quad \left(\rho_2\frac{\partial}{\partial\pi_2}\right),\quad \left(\rho_3\frac{\partial}{\partial\pi_3}\right),\quad \ldots,\quad \left(\rho_{n-1}\frac{\partial}{\partial\pi_{n-1}}\right),$$

where $\left(\rho_r\frac{\partial}{\partial\pi_r}\right)$ consists of $\binom{n}{r}$ terms

$$\Sigma\,\rho_{i_1 i_2\ldots i_r}\frac{\partial}{\partial\pi_{i_1 i_2\ldots i_r}},\qquad (\pi_{i_1 i_2\ldots i_r}=(xy\ldots)_{i_1 i_2\ldots i_r}),$$

the r suffixes taking values 1, 2, ..., n, and no two in one set being equal.

Example.—

If a_x^2, b_x^2 are two quadratics, and x, y, z, t denote four points, then the concomitant

$$(ab\mid xy)\,(ab\mid zt)$$

may be regarded as a polar of $(ab\mid xy)^2$, namely, it is $\frac{1}{2}\left(\rho_2\frac{\partial}{\partial\pi_2}\right)(ab\mid xy)^2$

where $\pi_2 = \overline{xy}$ and $\rho_2 = \overline{zt}$. In this example the polar consists of a single term. But more generally the polar of

$$(ab \mid xy)\,(cd \mid xy)$$

is

$$(ab \mid xy)\,(cd \mid zt) + (ab \mid zt)\,(cd \mid xy),$$

giving an example of a series of two terms derived by polarization.

Manifestly repeated operation with $\left(\rho_i \dfrac{\partial}{\partial \pi_i}\right)$, $\left(\sigma_i \dfrac{\partial}{\partial \pi_i}\right)$, &c., on a form F of order p_i in the variable π_i, produces in general a series of considerable complexity. Still more so if this can simultaneously go on for values of $i = 1, 2, \ldots, n - 1$. Our immediate aim is to express any single term of such a polar as an aggregate of polars of certain *standard forms* F together with the absolute concomitant which we shall denote by either π_n or E. Although at first sight this looks impossible, it can in fact be done, and is indeed important for the following reasons:

(1) Polarization is an invariant process (§8, p. 207, *example*).

(2) Any single term of a symbolic expression of any concomitant, whatever variables $x, y, \ldots, u, v, \ldots$ may be involved, is always a term of a polarized standard form F.

This last follows from the Fundamental Theorem. For if each $u, v \ldots$ which appears is treated as an $(n - 1)$th compound of the x's, then each symbolic factor of a term P either is free from variables or is an explicit linear function of an rth compound $(r = 1, 2, \ldots, n - 1)$, say ρ_r. As such it is a polar derived from π_r by the operator $\rho_r \dfrac{\partial}{\partial \pi_r}$. Hence P is certainly a term of a polar of a standard form.

Example.—

$(ab \mid zt)^2\,(\alpha\beta\gamma y)\,(cd \mid xy)$ is a term of

$$\left(\rho_2 \frac{\partial}{\partial \pi_2}\right)^2 \left(y \frac{\partial}{\partial x}\right) (ab \mid \pi_2)^2\,(\alpha\beta\gamma x)\,(cd \mid \pi_2).$$

3. Reduction to Standard Form.

The reduction of any form f to standard form depends upon two main ideas, one being the use of the Sylvester fundamental identity (§13, p. 93), and the other the theory of adjacent terms in a permanent (§1, p. 14).

We take the most general symbolic form as in §2, p. 198, involving cogredient symbols, and first consider a standard p-ic in one variable x,

$$f = a_{1x} a_{2x} \ldots a_{px}. \qquad \ldots\ldots \quad (5)$$

If we polarize this with regard to p cogredient variables x_1, $x_2, \ldots, x_p$ we obtain

$$\Delta f \equiv \left(x_1 \frac{\partial}{\partial x}\right) \left(x_2 \frac{\partial}{\partial x}\right) \ldots \left(x_p \frac{\partial}{\partial x}\right) f = \Sigma\, a_{1x_i} a_{2x_j} \ldots a_{px_k}, \quad (6)$$

summed for the p! permutations of the p suffixes $i, j, \ldots, k$. Here on the right is an example of a permanent (§1, p. 14), all the signs being positive. Let us call two of its terms adjacent if they differ by adjacent interchange of suffixes (§1, p. 14). Then any two terms T and T' can be connected by a series of adjacent terms $T_1, T_2, \ldots, T_\mu$, all belonging to Σ.

But the difference between two adjacent terms T_α, $T_{\alpha+1}$ leads to a Sylvester identity. In fact, if i and j are the two interchanged suffixes of the terms, we have

$$a_{qx_i} a_{rx_j} - a_{qx_j} a_{rx_i} = (a_q a_r \mid x_i x_j), \qquad \ldots \quad (7)$$

and the other $(p-2)$ factors of the terms are common. Thus

$$T_\alpha - T_{\alpha+1} = (a_q a_r \mid x_i x_j) \Pi a_{sx_k}, \qquad \ldots \quad (8)$$

where Π denotes the other $(p-2)$ unaffected factors. Hence by continued application to adjacent terms

$$\begin{aligned} T - T' &= (T - T_1) + (T_1 - T_2) + \ldots + (T_\mu - T') \\ &= \Sigma (a_q a_r \mid x_i x_j) \Pi a_{sx_k}, \qquad \ldots\ldots\ldots \quad (9) \end{aligned}$$

and by taking T' to be each term of the series (6) in turn and adding, we have

$$p!\ T - \Delta f = p!\ T - \Sigma T' = \Sigma\Sigma (a_q a_r \mid x_i x_j) \Pi a_{sx_k}. \quad (10)$$

This shows that, but for a non-zero numerical factor p!, any term T of Δf is equivalent to the polar Δf itself together with terms like those in the right member of (9) derived by convolution from Δf.

To simplify the notation let us provisionally write $\{x_i\}$ for any factor a_{qx_i}, $\{x_i x_j\}$ for any factor $(a_q a_r \mid x_i x_j)$, whatever a_q, a_r

may be, and so on for $\{x_i x_j x_k\}$, &c. We can now consider a standard form in two variables x, y of type

$$F = \Pi(a_q a_r \mid xy)\Pi a_x = \{xy\}\ \{xy\}\ \dots\ \{x\}\ \{x\}\ \dots, \qquad (11)$$

where there are j_2 factors $\{xy\}$ and j_1 factors $\{x\}$.

Any polar ΔF of this with regard to j_2 variables ρ_2, $\sigma_2 \dots$ leads to $j_2!$ terms such that the difference between adjacent terms gives a Sylvester identity, say

$$\begin{aligned}(a_q a_r \mid \rho_2)\,(a_s a_t \mid \sigma_2) - (a_q a_r \mid \sigma_2)\,(a_s a_t \mid \rho_2)& \\ = \Sigma\,(\dots \mid x_i x_j x_k)\,(\dots \mid x_l) + \Sigma\,(\dots \mid x_i x_j x_k x_l),& \qquad (12)\end{aligned}$$

where $\rho_2 = \overline{x_i x_j}$, $\sigma_2 = \overline{x_k x_l}$.

Hence, arguing as before, we deduce that any term of type

$$\{x_i x_j\}\{x_k x_l\}\ \dots\ \{x\}\{x\}\ \dots \qquad . \quad . \quad (13)$$

is equivalent to a polar of a standard form (11), together with forms with more than two x's convolved in one factor. These last may introduce a factor $\{x_l\}$ as in the first term on the right of (12); and this, along with the factors $\{x\}\ \{x\}\ \dots$ of (11), can be dealt with as a polar with regard to $\left(x_l \frac{\partial}{\partial x}\right)$, leading, as in (10), to a new factor $\{x_l x\}$, for which the argument may be repeated.

Combining these results we gather that any term

$$\{x_i x_j\}\{x_k x_l\}\ \dots\ \{x_m\}\{x_q\}\ \dots \qquad . \quad . \quad (14)$$

is equivalent to a polar of (11), together with forms involving three or more x's convolved.

Proceeding in this way and using the Sylvester identity for each further case in turn, with $r = 3, 4 \dots$, we arrive finally when $r = n - 1$ at the case where n variables $x, y, \dots, t$, and standard form F are given, such that

$$F = \Pi\{xy\dots st\}\ \Pi\ \{xy\dots s\}\dots\Pi\ \{xy\}\ \Pi\ \{x\}, \quad (15)$$

and *any form* f *involving any number of variables* x_1, ρ_2, $\sigma_2 \dots$, ρ_r, $\sigma_r \dots$, ρ_{n-1}, σ_{n-1} *is expressible as a series of terms* ΔF *derived from such forms* F *by polarization.*

Here, the first factor Π gives the absolute concomitant of the field and all the other factors involve at most $(n-1)$ variables $x, y, \dots, s$. This establishes the theorem of Clebsch.

In this final formula (15) each factor of the products has all its variables x, y . . . explicitly stated. The other symbols are implicit. They are the symbols a_1, a_2 . . . of the original form f in some order or other. Inasmuch as the process of this reduction of f to $\Sigma\Delta F$ is entirely composed of repetitions of the Sylvester identity, which preserves the symbols but only deranges them, it follows that any symbols convolved in the original form f are still convolved, implicitly or explicitly, in the standard form F.

Corollary.—*By taking the dual co-ordinates* $\mathrm{p_i} = \pi_{\mathrm{n-i}}$ *we can throw any standard form into a product of bracket factors*

$$F = \Pi(uv\ldots wu')\,\Pi(ab\ldots u)\,\Pi(a'b'\ldots uv)\ldots\Pi(a''uv\ldots w). \quad (16)$$

There are in fact four ways of writing each factor of F. Thus

$$\begin{aligned}(ab\ldots d \mid xy\ldots z) &= (ab\ldots d \mid \pi_{n-i}) \\ &= (ab\ldots dp_i) = (ab\ldots duv\ldots). \quad (17)\end{aligned}$$

In practice the process of deriving the standard forms for a given expression $f(x_1, x_2 \ldots)$ is exceedingly complicated except in the simplest cases. The present treatment follows the algebraic method as used for binary forms.[1] The usual treatment follows the methods of Capelli who bases all on differential operations rather than algebraic permutations.

4. The Gordan-Capelli Series.

Let us apply the preceding methods to the form

$$f = a_x^p\, b_y^q\, c_z^r \ldots d_t^s, \quad . \quad . \quad . \quad . \quad . \quad . \quad (18)$$

where there are $k \geqslant 2$ cogredient sets $x, y, z, \ldots, t$. First suppose $k < n$. Then such a form f is a term of a polar of

$$a_x^p b_x^q c_x^r \ldots d_x^s,$$

and, treated as above, f is equal to a sum of terms where the most advanced convolution of the variables is

$$\{xyz\ldots t\} \equiv (abc\ldots d \mid xyz\ldots t) \equiv K; \quad . \quad (19)$$

for all the k letters before or after the vertical line must differ, so that there is just this one possibility.

---

[1] Grace and Young, *Algebra of Invariants* (1903), 42–46.

Terms which do not contain K are due to polars of standard forms F_0 involving $k-1$ variables, say all but x. So we write

$$f = \Sigma\,\Delta_0 F_0 + K\phi, \quad . \quad . \quad . \quad . \quad (20)$$

Δ_0 denoting an aggregate of polar operations, and ϕ necessarily containing x to degree $p-1$, y to degree $q-1$, and so on. Treating ϕ in the same way as f we have

$$\phi = \Sigma\,\Delta_1 F_1 + K\psi$$

where $\Delta_1 F_1$ has the same general meaning as $\Delta_0 F_0$, and ψ is of degrees $p-2$, $q-2$, . . . , in the variables.

Proceeding in this way we exhaust one of the variables in h steps where h is the least of the exponents $p, q, \ldots, s$; and thereby we obtain the *Gordan-Capelli Series*,

$$f = \Sigma\,\Delta_0 F_0 + K\,\Sigma\,\Delta_1 F_1 + K^2\,\Sigma\,\Delta_2 F_2 + \ldots + K^h\,\Delta_h F_h. \quad (21)$$

Here each F_i is a form involving at most $(k-1)$ different sets $y, z, \ldots, t$; K is the kth compound inner product $(ab \ldots d \mid xy \ldots t)$; and Δ_i is a polar operation. Some of the coefficients of powers of K may in particular cases be zero.

Secondly, if $k = n$, the expression K is replaced by an actual product

$$(ab \ldots d)\,(xy \ldots t)$$

involving the absolute invariant $(xy \ \ldots \ t) = E$ of the field. In this case the Gordan-Capelli series is

$$\begin{aligned} f = \Sigma\,\Delta_0 F_0 &+ (xy \ldots t)\Sigma\,\Delta_1 F_1 \\ &+ (xy \ldots t)^2 \Sigma\,\Delta_2 F_2 + \ldots + (xy \ldots t)^h \Delta_h F_h \end{aligned} \qquad (22)$$

Here the coefficients of the series are polars of forms F_i involving at most $(n-1)$ sets $y, z, \ldots, t$.

Thirdly, if $k > n$ no corresponding form K exists, so that we have

$$f = \Sigma\,\Delta_0 F_0 \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (23)$$

expressing f as a series of polars of forms involving at most $(n-1)$ sets $y, z, \ldots, t$.

Various alternative expressions of K are furnished by (17).

In particular, if $k = n-1$, we write u instead of p_1, where the set u denotes a prime, such that

$$u_1 = +\,(xy \ldots t)_{23 \ldots n}, \ \&c.,$$

and the series now takes the form

$$f = \Sigma \Delta_0 F_0 + (ab \ldots du) \Sigma \Delta_1 F_1 + (ab \ldots du)^2 \Sigma \Delta_2 F_2 + \ldots\ldots$$

5. Examples of the Series for Binary and Ternary Fields.

In the binary field for a form

$$f = a_x^m b_y^n \qquad (n \leqslant m) \qquad \ldots\ldots \qquad (24)$$

with two sets x and y the series was originally given by Clebsch and Gordan

$$\underset{k=0}{-} \frac{\binom{m}{k}\binom{n}{k}}{\binom{m+n-k+1}{k}} \frac{(m-k)!}{(m+n-2k)!} (xy)^k D_{xy}^{\,n-k} \{(ab)^k a_x^{m-k} b_x^{n-k}\} \qquad (25)$$

where D_{xy} denotes the polar operator $\left(y\frac{\partial}{\partial x}\right)$. Cf. Grace and Young, *Invariants*, p. 55.

In this case the coefficient of $(xy)^k$ is the $(n-k)$th polar of a form

$$f_k = (ab)^k a_x^{m-k} b_x^{n-k}$$

depending on only one set of variables.

In the ternary field where now $a_x = a_1 x_1 + a_2 x_2 + a_3 x_3$, the corresponding series for $f = a_x^m b_y^n (n \leqslant m)$ is

$$f = \sum_{k=0} \frac{\binom{m}{k}\binom{n}{k}}{\binom{m+n-k+1}{k}} \frac{(m-k)!}{(m+n-2k)!} (abu)^k D_{xy}^{\,n-k} \{a_x^{m-k} b_x^{n-k}\}$$

where $u_1 = (xy)_{23}$, &c. And more generally the series for

$$f = a_x^m b_y^n c_z^p$$

is

$$f = \Sigma \Delta_0 F_0 + (xyz) \Sigma \Delta_1 F_1 + (xyz)^2 \Sigma \Delta_2 F_2 + \ldots.$$

6.[1] Normal Forms.

Ground forms which can be symbolized as

$$(abc \mid xyz)^i (ab \mid xy)^j (a \mid x)^k \qquad i, j, k \geqslant 0, \quad . \quad (26)$$

with symbols as well as variables appearing in the characteristic *standard* order are called normal forms. In this example we assume $n > 3$, to prevent the first factor from reducing.

It is obvious that for a normal form any invariant linear in its coefficients must vanish. For every outer product $(abc \ldots)$ must contain a repeated symbol.

Again, by polarization with $\left(x\frac{\partial}{\partial y}\right)$, $\left(x\frac{\partial}{\partial z}\right)$, or $\left(y\frac{\partial}{\partial z}\right)$ we manifestly obtain a zero result. This leads to an important theorem:

---

[1] This section may be omitted on a first reading.

If, when $n > 4$, *the form* $f = a_x^p b_y^q c_z^r d_t^s$, $p \geqslant q \geqslant r \geqslant s$, *satisfies the three differential equations*

$$D_{tz}\phi = 0, \qquad D_{zy}\phi = 0, \qquad D_{yx}\phi = 0, \qquad . \quad (27)$$

it can be written in the alternative normal form

$$C(abcd \mid xyzt)^s (abc \mid xyz)^{r-s} (ab \mid xy)^{q-r} (a \mid x)^{p-q}$$

where C *is a numerical constant, and* $D_{yx} = \left(x \dfrac{\partial}{\partial y}\right)$, *&c.*

Proof.—

The argument will hold for any number of variables. Consider the matrix of polar operations

$$D = \begin{bmatrix} t_t & z_t & y_t & x_t \\ t_z & z_z & y_z & x_z \\ t_y & z_y & y_y & x_y \\ t_x & z_x & y_x & x_x \end{bmatrix} = [\beta_\alpha] \qquad (28)$$

where the typical element β_α denotes a polar operation, e.g. $t_z = \left(t \dfrac{\partial}{\partial z}\right) = D_{zt}$.

Since $\beta_\alpha = \beta_\gamma \gamma_\alpha - \gamma_\alpha \beta_\gamma$ holds for any three α, β, γ among x, y, z, t, we infer that all elements above and to the right of the diagonal z_t, y_z, x_y are expressible in terms of these. Hence, for reasons used in (23), p. 116, the equations (27), which we can call the diagonal equations, imply a whole triangle of equations. This is true in general for k variables $x, y, \ldots, s, t$.

Again, if H is the Capelli operator (25), p. 117, answering to the matrix D, its expansion as a determinant consequently loses all terms except the leading diagonal

$$t_t(z_z + 1)(y_y + 2)(x_x + 3)$$

when it operates on f, if all these elements of the upper triangle are zero.

But by Euler's theorem for homogeneous functions

$$x_x f = pf, \qquad y_y f = qf, \qquad z_z f = rf, \qquad t_t f = sf.$$

Hence $$Hf = (p + 3)(q + 2)(r + 1)sf.$$

Once more, writing $H = \Sigma (xyzt)_{ijkl} \left(\frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\ \frac{\partial}{\partial z}\ \frac{\partial}{\partial t}\right)_{ijkl}$ and operating directly on f, we obtain

$$pqrs\,(abcd \mid xyzt)\, a_x^{p-1} b_y^{q-1} c_z^{r-1} d_t^{s-1}$$

Accordingly

$$(p+3)(q+2)(r+1)\,sf = pqrs\,Kf_1, \qquad (29)$$

where

$$K = (abcd \mid xyzt), \qquad f_1 = a_x^{p-1} b_y^{q-1} c_z^{r-1} d_t^{s-1}.$$

Let the polar operator β_α ($\alpha \neq \beta$) denote any element of (28), not on the leading diagonal. Then by actual differentiation we find,

$$\beta_\alpha K f_1 = (\beta_\alpha K) f_1 + K \beta_\alpha f_1 = K \beta_\alpha f_1,$$

so that K *commutes* with the operator β_α, when acting on f_1. Hence $K\beta_\alpha f_1 = 0$, whenever $\beta_\alpha K f_1$ and therefore, by (29), $\beta_\alpha f$ vanishes. So that Kf_1 satisfies exactly the same relations (27) as f, provided that, for the purpose of differentiation, K is regarded as a constant.

Then if $s > 1$, we deduce, similarly to (29), that $Kf_1 = Kc_1 Kf_2$, and therefore

$$f = CK^2 f_2, \quad f_2 = a_x^{p-2} b_y^{q-2} c_z^{r-2} d_t^{s-2},$$

where c_1, C are numerical non-zero constants.

Similarly s such operations lead to

$$f = C_1 (abcd \mid xyzt)\ a_x^{p-s} b_y^{q-s} c_z^{r-s},$$

where C_1 is numerical, and non-zero.

Finally by $(r-s)$ operations with $\left(xyz \,\middle|\, \frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\ \frac{\partial}{\partial z}\right)$ followed by $(q-r)$ with $\left(xy \,\middle|\, \frac{\partial}{\partial x}\ \frac{\partial}{\partial y}\right)$ we obtain the desired normal form, so proving the theorem.

EXAMPLES

1. The necessary and sufficient conditions for f to be symbolized by a perfect pth power

$$f = (abc \ldots e \mid xy \ldots t)^p$$

are that it satisfies $2(k-1)$ conditions

$$D_{xy} f = D_{yz} f = \ldots = D_{st} f = 0, \quad D_{ts} f = \ldots = D_{zy} f = D_{yx} f = 0$$

among the k variables $x, y, \ldots, s, t$.

2. What are the requisite conditions for a function of n sets of variables $x, y, \ldots, t$ to be a perfect pth power of the determinant $(xy \ldots t)$?

[Put $k = n$ in Ex. **1**.

3. The necessary and sufficient conditions for a symbolic form in two sets x, y to be a perfect pth power $(ab \mid xy)^p$ are that it satisfies two differential equations

$$\left(y\frac{\partial}{\partial x}\right)f = 0, \quad \left(x\frac{\partial}{\partial y}\right)f = 0.$$

4. A quaternary line complex is a form in six variables, $p_{12}, p_{13}, p_{14}, p_{23}, p_{34}, p_{42}$, where $p_{ij} = (xy)_{ij}$. Prove that it can always be symbolized in the normal form $(abp)^n \equiv (ab \mid xy)^n$.

[Use the differential equations of Ex. 3 on the non-symbolic form.

7. Historical Note.

The results given in this chapter cover a long period of study. The Gordan-Capelli binary series was first given by Clebsch and Gordan,[1] and next it was extended to the general case by Clebsch and Capelli.[2]

These normal forms of §6 were called *primary covariants* by Deruyts (*Essai*, . . . (1891)), who also studied this general problem, although the theory goes back to Clebsch[3], Gordan, Mertens,[4] and Study.[5] The general theory is given by E. Noether[6] who uses the theorem of corresponding matrices, and by Weitzenböck[7] who introduces complex symbols. A purely algebraic discussion free from differential operators can be based upon the far-reaching results of Frobenius[8] and A. Young.[9]

---

[1] *Math. Annalen*, **5** (1872), 95–122.

[2] For ternary forms, *Battaglini*, **18** (1880). For n-ary forms, *Mem. del. R. Acc. dei Lincei* (1882), (1891), (1892), and *Math. Annalen*, **27** (1885). See above, p. 254 (21).

[3] *Göttinger Nachrichten*, **17** (1872). Ternary and general.

[4] *Wiener Berichte*, **98** (1899). Quaternary.

[5] *Methoden*, p. 54. Ternary.

[6] *Math. Annalen*, **77** (1915), 93: *Crelle*, **139** (1910), 118 *seqq.*

[7] *Invariantentheorie* (1923), V, pp. 121–159.

[8] *Berliner Sitzungsberichte*, **1** (1897); **2** (1899).

[9] *Algebra of Invariants*, Chapter XVI: *Proc. London Math. Soc.*, **33** (1901) and **34** (1903), (2) **28** (1928).

CHAPTER XVII

APPLICATIONS OF CLEBSCH'S THEOREM. APOLARITY AND CANONICAL FORMS

1. Similar Forms.

When a number of forms

$$f_1, f_2, \ldots, f_N, \ldots \qquad (1)$$

all have the same sets of variables and are all of the same respective orders $[p, q, \ldots]$ in these variables, they are called *similar forms.* For example, we may have a system of ternary quadratics, in which case $n = 3$, $p = 2$, and one set x of variables is used.

Let N be the number of the coefficients

$$A_1, A_2, \ldots, A_N \qquad (2)$$

in f_1 and therefore in each similar form. For ternary quadratics this N would be six. Further let $A, B, \ldots, G, H$ denote these coefficient sets of $N + 1$ similar forms $f_1, f_2, \ldots, f_N, f_{N+1}$.

From these we can build a vanishing determinant

$$\Delta = \begin{vmatrix} A_1 & A_2 & \ldots & A_N & f_1 \\ B_1 & B_2 & \ldots & B_N & f_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ G_1 & G_2 & \ldots & G_N & f_N \\ H_1 & H_2 & \ldots & H_N & f_{N+1} \end{vmatrix} = 0, \quad . \quad . \quad (3)$$

because the last is a linear function of the other N columns. Let the expansion of this by its last column be

$$K_1 f_1 + K_2 f_2 + \ldots + K_{N+1} f_{N+1} = 0. \quad . \quad . \quad (4)$$

This shows that, as in §1, p. 73,

Any N + 1 *similar forms are linearly dependent.*

If the rank r of the matrix of the first N columns is less than N, then the forms are not the most general, but can be expressed in terms of r forms suitably chosen. If $r = 2$ each form can be written as

$$\lambda f + \lambda' f'$$

in terms of two linearly independent similar forms. Then they are said to make a *pencil* of forms. For $r = 3$ they make a *net*

$$\lambda f + \lambda' f' + \lambda'' f''.$$

The fundamental property of similar forms is this:

Each set A, B, . . . *behaves like a cogredient linear form in the field of category* N.

For if $A \rightarrow A'$, $B \rightarrow B'$, . . . denote the transformations induced by that of the variables $x \rightarrow x'$, then the coefficient matrices of these transformations are all precisely the same because the sets A, B, . . . are similar. It follows that the co-factor of f_{N+1} in Δ, say

$$K_{N+1} = | A_1 B_2 \ldots G_N |,$$

is an invariant, since it is a typical bracket factor for the field of category N. So also are each of K_1, K_2, ...

2. Types.

We already know that the Aronhold operator (§10, p. 141)

$$\left(B \frac{\partial}{\partial A}\right) = \sum_{i=1}^{N} B_i \frac{\partial}{\partial A_i}$$

produces a concomitant when it operates on a concomitant involving A. Indeed the process is analogous to the polar process involving variables, and thereby it leads to a theorem, first given by Peano, which will be seen to play for the coefficients exactly the same part that the theorem of Clebsch played for the variables.

Let all such Aronhold operators involving similar sets of coefficients A, B, C, . . . be utilized, and a rational integral combination of these acting on a concomitant be called an *Aronhold process*. Then every concomitant so derivable from one and the same original concomitant is said to be of the same *type*.

EXAMPLES

1. Any concomitant can be rendered multilinear (§10, p. 141) in its ground form coefficients, by Aronhold processes. All concomitants of the same type can be brought to the same multilinear form.

2. An Aronhold operator $\Sigma a_i a_j \dots \dfrac{\partial}{\partial A_{ij\dots}}$ acting on a form linear in the coefficients $A_{ij\dots}$ automatically replaces the actual by the symbolic coefficients. One such operator for each coefficient set reduces the form entirely to symbols.

3. Ternary cubics a_x^3, b_x^3, c_x^3, d_x^3, e_x^3 have an invariant of type $(abc)^3$. For each of $(abd)^3$, $(bce)^3$, . . . can be derived by an Aronhold process.

3. Peano's Theorem.

It is obvious from (4) that if one of Δ_i is non-zero, say K_{N+1}, then any form can be expressed in terms of the first N similar forms. Each coefficient of f_{N+1} is given as a *rational* function of those of f_1, . . . , f_N, with K_{N+1} appearing in the denominator. If, further, the irreducible system, given by Gordan's theorem, is known for the N forms, a *rationally* complete system for any extra simultaneous similar forms naturally follows. But we can go further and find an *integrally* complete system in general, once we know it for $N-1$ similar forms: and this brings us to the theorem.[1]

Peano's Theorem.—*With the possible exception of the determinant* K, *linear in the coefficients of* N *similar forms each with* N *coefficients, every polynomial concomitant of any number of such forms is expressible by the complete system of* N — 1 *such forms, and by types derived from this system.*

Proof.—

We regard the concomitant as a polynomial, homogeneous in each set of N coefficients A. Selecting the coefficient sets of the first N of $N+i$ such forms, we express every polynomial concomitant as a Gordan-Capelli series (p. 254 (21))

$$\Delta_0\phi_0 + K\Delta_1\phi_1 + K^2\Delta_2\phi_2 + \dots,$$

where K is the determinant K_{N+1} in these N forms, ϕ_i is a function of at most $N-1$ sets of coefficients, and Δ_i is an Aronhold process.

---

[1] *Atti di Torino*, **17** (1881), p. 580; D. Hilbert, *Schwarz-Festschrift* (1914); E. Noether, *Math. Annalen*, **77** (1915), p. 93; for binary forms, see Grace and Young, pp. 321, 349, 358; Weitzenböck, *Invariententheorie*, p. 162.

By the mode of constructing such a series, each ϕ_i is invariantive since we started with an actual concomitant. Thus the series is entirely composed of types belonging to $N-1$ ground forms at most, together with K, involving N ground forms linearly. Q.E.D.

In some cases K itself is reducible, as in the binary forms of odd order.[1] In some it is certainly irreducible as in the case of six conics.

Example.—

In the quaternary field, a complete study of concomitants is effected by confining oneself to ground forms in three types of variable $x = \overline{uvw}$, $p = \overline{uv}$, and u, together with polar forms, while the knowledge of all possible types of concomitant for a given type of ground form, say a quadratic in x which has ten terms $\Sigma a_{ij} x_i x_j$, is complete if we know those of nine quadrics together with the ten-rowed determinant linear in the coefficients of ten quadrics.

4. Dual Similar Forms.

Just as there are dual systems of variables

$$\begin{array}{ll} & x, y, \ldots, \pi_2, \ldots, \pi_3, \ldots, \pi_{n-1} \\ \text{and} & u, v, \ldots, p_2, \ldots, p_3, \ldots, p_{n-1}, \ldots \end{array} \qquad (5)$$

so we may consider certain forms f and ϕ to be dual of each other. Symbolically we merely have to interchange italic and Greek letters $a, b, c, \alpha, \beta, \gamma, \ldots$ throughout.

For example, a_x^p, u_α^p are dual forms of order p in the original variables; $(ab \mid xy)^p$, $(\alpha\beta \mid uv)^p$ are dual forms in second compound variables. More generally

$$\left.\begin{array}{l} f = a_x a'_y \ldots (bc \mid \pi_2) \ldots (def \mid \pi_3) \ \ldots \\ \phi = u_\alpha v_{\alpha'} \ldots (\beta\gamma \mid p_2) \ldots (\delta\epsilon\zeta \mid p_3) \ldots \end{array}\right\} \quad . \quad . \quad (6)$$

are called *similar dual forms* when they possess corresponding symbolic factors. Manifestly they have the same orders in their corresponding variables (5), and the same number N of actual coefficient sets, say

$$\left.\begin{array}{l} A = [A_1, A_2, \ldots, A_N] \\ R = \{\rho_1, \rho_2 \ \ldots, \rho_N\} \end{array}\right\} . \quad . \quad . \quad . \quad (7)$$

---

[1] Grace and Young, *loc. cit.*

Obviously the preceding results of this chapter apply to a system of similar forms $\phi_1, \phi_2, \dots, \phi_N, \dots$.

In general the sets R and A are arbitrary and independent, for this duality merely refers to their structure.

Dual similar forms $f = a_x^p$, $\phi = u_\alpha^p$ *possess a simultaneous invariant linear in each, namely*

$$(f, \phi)_N = a_\alpha^p = c_1 A_1 \rho_1 + c_2 A_2 \rho_2 + \dots + c_N A_N \rho_N,$$

where each c_i is the multinomial coefficient occuring in the ith term of both f and ϕ.

This invariant can be generated non-symbolically by the operator

$$Q = \left(\frac{\partial}{\partial x} \middle| \frac{\partial}{\partial u}\right) = \frac{\partial}{\partial x_1}\,\frac{\partial}{\partial u_1} + \dots + \frac{\partial}{\partial x_n}\,\frac{\partial}{\partial u_n}.$$

In fact

$$Qf\phi = \Sigma \frac{\partial}{\partial x_i}\,\frac{\partial}{\partial u_i} f\phi = \Sigma \frac{\partial f}{\partial x_i}\,\frac{\partial \phi}{\partial u_i} = p^2 \Sigma\, a_x^{p-1} a_i u_x^{p-1} a_i$$
$$= p^2 a_\alpha a_x^{p-1} u_\alpha^{p-1}.$$

Hence $$Q^p f\phi = p!\,p!\,a_\alpha^p = p!\,p!\,(f, \phi)_N.$$

EXAMPLES

1. The ternary quadratics $ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy$ and $Au^2 + Bv^2 + Cw^2 + 2Fvw + 2Gwu + 2Huv$ have a bilinear invariant $\Theta = aA + bB + cC + 2fF + 2gG + 2hH$.

2. Adapt this theorem to *binary forms.*

If $f = a_x^p$ is a binary p-ic $(a_0, a_1, \dots, a_p \between x_1, x_2)^p$, a dual form is also a binary p-ic but its coefficients are reversed with alternate signs changed. Thus

$$a_0, a_1, a_2, a_3 \qquad b_3, -b_2, b_1, -b_0$$

are dual sets for binary cubics. For by (49), p. 145, contragredient *binary* variables u_1, u_2 are cogredient with $x_2, -x_1$, so that the dual form u_β^p is a polarized form of, say, b_x^p, the polar operator being $\left(u_2 \frac{\partial}{\partial x_1} - u_1 \frac{\partial}{\partial x_2}\right)^p$.

3. For binary forms a_x^p, b_x^p this invariant is their pth transvectant $(ab)^p = a_0 b_p - p a_1 b_{p-1} + \binom{p}{2} a_2 b_{p-2} - \dots + (-)^p a_p b_0$.

4. Two general similar forms $a_x b_y \dots$, $u_\alpha v_\beta \dots$ have a corresponding invariant $a_\alpha b_\beta \dots$ derived by a product of operators $\left(\frac{\partial}{\partial x} \middle| \frac{\partial}{\partial u}\right)$, $\left(\frac{\partial}{\partial y} \middle| \frac{\partial}{\partial v}\right)$, ...

5. More generally, two dual similar forms involving compounds π_i, p_i have such an invariant, derived from a product of operators $\left(\frac{\partial}{\partial \pi_i} \,\middle|\, \frac{\partial}{\partial p_i}\right)$.

6. Derive $a_\alpha (bc \,|\, \beta\gamma)^2$ from $a_x (bc \,|\, \pi_2)^2$, $u_\alpha (\beta\gamma \,|\, p_2)^2$: and write down a typical term in non-symbolic notation.

5. Apolarity.

Definition.—*Two dual similar forms* f, ϕ *are apolar if their lineo-linear invariant* $(\mathrm{f}, \phi)_N$ *vanishes.*

When two forms f, ϕ have this property very many interesting geometrical facts can be deduced. We shall confine ourselves to one aspect of the case, namely to the discovery of the so-called canonical forms. But as a preliminary to this, a few properties of apolar forms are useful.

(i) First, there is one dual form ϕ apolar to each of $N-1$ given linearly independent similar forms $f_1, f_2, \dots, f_{N-1}$; for this amounts to giving $N-1$ linear equations

$$c_1 A_1 \rho_1 + c_2 A_2 \rho_2 + \dots + c_N A_N \rho_N = 0, \quad . \quad . \quad (8)$$

one for each of the $N-1$ sets of coefficients $A, B, \dots, F$. The c's are the same throughout, and the equations determine the set ρ, and therefore the form ϕ, to a constant factor.

(ii) Next, if r of the coefficients A happen to vanish and the complementary $(N-r)$ coefficients of ϕ vanish, then (8) is satisfied, so that f and ϕ are apolar.

(iii) Again, if ϕ is apolar to each of $f_1, f_2, \dots$ it is also apolar to any linear combination of $f_1, f_2, \dots$. Further a linear combination of r forms f can be apolar to such a combination of $N-r$ forms ϕ. These results all follow from the condition (8).

6. Apolarity of Dissimilar Forms.

When forms f, f', possess the same variable sets but to different orders, they may be reduced to a common order by polarization. For simplicity let us deal with *one* variable set x only. Then if $f = a_x^p, f' = b_x^q$, $q > p$, we polarize the latter $(q-p)$ times with an *arbitrary* cogredient set y and obtain the form $f'' = b_y^{q-p} b_x^p$ similar to f as regards x.

Let M be the number of separate terms in the non-symbolic expression of f'' as a polynomial in the arbitrary set y. Then f'' is in effect a set of M linearly independent p-ics in x. Now if $N > M$ we can find $N - M$ forms $u_\alpha{}^p$, apolar to each such portion of f''. We therefore take this as the definition of apolarity of f' and ϕ when the orders differ. Thus, in symbolic notation, we infer that

Two forms $\mathrm{f}' = \mathrm{b}_\mathrm{x}{}^\mathrm{q}$, $\phi = \mathrm{u}_\alpha{}^\mathrm{p}$ $(\mathrm{q} > \mathrm{p})$ *are apolar if the covariant* $\mathrm{b}_\alpha{}^\mathrm{p}\,\mathrm{b}_\mathrm{y}{}^{\mathrm{q}-\mathrm{p}}$ *vanishes identically for all values of* y.

Dually, *if* $\mathrm{q} < \mathrm{p}$, f′ *and* ϕ *are apolar when the contravariant* $\mathrm{b}_\alpha{}^\mathrm{q}\,\mathrm{v}_\alpha{}^{\mathrm{p}-\mathrm{q}}$ *vanishes identically for all values of* v.

EXAMPLES

1. Let $f = a_x{}^3$, $\varphi = u_\alpha{}^2$ be a ternary cubic and contravariant quadratic. Then $a_x{}^2\, a_y$ has 3 terms in y, while $u_\alpha{}^2$ has 6 terms: $M = 3$, $N = 6$. Hence three linearly independent conics exist which are apolar to f.

2. A binary quartic $a_x{}^4$ has two terms in y among its cubic polars $a_x{}^3\, a_y$: $M = 2$, $N = 5$.

3. If $f = a_x{}^p$, $f' = b_x{}^q$, $\varphi = u_\alpha{}^{p+q}$, then ff' is apolar to φ if either f or f' is apolar to φ.

4. The ternary p-ic $a_x{}^p$ is apolar to the pth power of a linear form u_y if $a_y{}^p = 0$.

Geometrically, if a curve of order p passes through a point y, the point reckoned p times is a dual apolar form.

5. If $a_x{}^p$ is apolar to the $(p-1)$th power of a linear form u_y then $a_y{}^{p-1}\, a_z = 0$ for all values of z. Geometrically, what is the point y?

[A double point on the curve.

6. If $a_x{}^p$ is apolar to the $(p-1)$th powers of k different linear forms the geometrical locus $a_x{}^p = 0$ has k distinct double points.

This is true for all fields, ternary and higher.

7. Canonical Forms.

Let

$$A_1, A_2, \ldots, A_N \quad \text{and} \quad A_1', A_2', \ldots, A_N' \qquad (9)$$

be the coefficients of a form f before and after linear transformation of its variables, so that each A_i' is a linear function of the A's. Thus

$$A_i' = \theta_{i1}A_1 + \theta_{i2}A_2 + \ldots + \theta_{iN}A_N, \quad . \quad . \quad (10)$$

where each θ_{ij} is a function of the elements $\xi_1, \ldots, \omega_n$ in the matrix M of the transformation.

Then a very pertinent question arises, how far can these A''s be arbitrarily assigned? In particular can some, and if so how many, of them be zero? Can we, in fact, throw f into a simpler form by a suitable transformation?

Answers to these questions are ready to hand. Thus if capital letters X, Y, Z, X_i denote linear functions of the variables, and in each case the *general* form is under discussion, then:

1. *The binary cubic*[1] *can be written in the canonical form* $X^3 + Y^3$.
2. *The binary quintic*, $X^5 + Y^5 + Z^5$.
3. *The binary quartic*, $X^4 + 6mX^2Y^2 + Y^4$.
4. *The ternary cubic*, $X^3 + Y^3 + Z^3 + 6mXYZ$.
5. *The ternary quartic*, $S_2S_3 - S_1^2$, *in terms of three ternary quadratics.*
6. *The quaternary cubic*, $X_1^3 + X_2^3 + X_3^3 + X_4^3 + X_5^3$.
7. *The ternary quartic* cannot *be written as the sum of five fourth powers of linear forms.*

Let us typify these problems by stating each in the form

$$f(x) \equiv F(X) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (11)$$

(or $F(S)$ in case (5)). Each X is a linear function of the original variables x_i; the total number of terms on the left side is N; and this is an identity for all values of x_i. Hence N separate equations connect the coefficients of terms in x on the left and right.

Let $X = a_1x_1 + a_2x_2 + \ldots + a_nx_n$, $Y = b_x$, $Z = c_x$, and so on. Here we have n undetermined coefficients for each X or Y or Z, giving for cases (1), (2), (6) above $2 + 2$, $2 + 2 + 2$, $4 + 4 + 4 + 4 + 4$ such unknowns. Now these exactly tally with N the number of terms in the given form $f(x)$.

Thus the binary cubic has four terms, and $X^3 + Y^3$ written as $(a_1x_1 + a_2x_2)^3 + (b_1x_1 + b_2x_2)^3$ has four unknowns a_1, a_2, b_1, b_2. In general we can solve these N equations for N unknowns and obtain a finite number of sets of values a_i, b_i, . . . which reduce $f(x)$ to $F(x)$. This is called the test by counting constants. We call all these unknowns, together with further coefficients m in the canonical form, the *parameters*. Parameters occurring in X, Y, Z, &c., are *implicit*; others such as m are *explicit*.

---

[1] Note that the form $X^3 + Y^3$ is inadmissible if the cubic has a repeated factor.

Examples.—

Case (3): $N = 5 = 2 + 2 + 1$. Here m is the extra unknown.

Case (5): $N = 15 < 6 + 6 + 6$. Each S has six unknown coefficients. Presumably three can be arbitrary.

Case (7): $N = 15 = 3 + 3 + 3 + 3 + 3$. *This passes the test.*

8. Counting Constants is not Sufficient.

Historically the ternary quartic, in case (7), provides the key to what follows, because at one time it was assumed to fall in with the general law. But an easier example of the inadequacy of this test is furnished by attempting to write the binary cubic $X^3 + X^2Y$. Here $2 + 2 = 4$ satisfies the test, but it is insufficient because $F(X)$ contains a repeated factor whereas the original cubic need not. Something more is required; and it is supplied by the Lasker-Wakeford theorem.[1]

The Lasker-Wakeford Theorem.

A form F(X, m) *which contains not less than* N *parameters among the* k *auxiliary forms* X *and* r *explicit coefficients* m, *is, or is not, a legitimate canonical form of* f(x) *provided there is not, or there always is, a form* ϕ(u), *dual to* f(x) *and apolar to each of the* k + r *derivatives* ∂F(X, m) / ∂X, ∂F(X, m) / ∂m.

The forms X *need not necessarily be linear.*

Before proving this paradoxical and very curious theorem, let us illustrate its scope in cases (3) and (5). For the binary quartic

$$\tfrac{1}{4}\frac{\partial F}{\partial X} = X^3 + 3mXY^2, \quad \tfrac{1}{4}\frac{\partial F}{\partial Y} = 3mX^2Y + Y^3, \quad \tfrac{1}{6}\frac{\partial F}{\partial m} = X^2Y^2.$$

Treat X, Y as variables, and U, V as duals.

Now if $\phi = B_0U^4 + B_1U^3V + B_2U^2V^2 + B_3UV^3 + B_4V^4$ is a quartic apolar to X^2Y^2, then (p. 263, Ex. **3**) $B_2 = 0$. If it is also apolar to a cubic X^3 (whose coefficients are 1, 0, 0, 0) then each first polar $\left(U'\dfrac{\partial}{\partial U} + V'\dfrac{\partial}{\partial V}\right)\phi$ is apolar (p. 265, Ex. **3**). Hence the apolar condition is $4U'B_0 + V'B_1 = 0$, so that

---

[1] E. Lasker, *Math. Annalen*, **58** (1904), 434–440. E. K. Wakeford, *Proc. London Math. Soc.*, **2**, **18** (1918–19), 403–410.

B_0, B_1 both vanish. If ϕ also is apolar to Y^3, then $B_3 = B_4 = 0$. Hence an apolar form ϕ does not always exist, for it disappears in the case when $m = 0$. By the theorem this is enough to prove $X^4 + 6mX^2Y^2 + Y^4$ a legitimate form.

EXAMPLE

Prove φ non-existent for the cubic $X^3 + Y^3$ but existent for X^2Y.

9. Proof of the Theorem.

Let the parameters ν in number be $l_1, l_2, \ldots, l_\nu$, so that the assumed identity (11) leads to N ($\leqslant \nu$) equations of the type

$$A_i = f_i(l_1, l_2, \ldots, l_\nu) \quad . \quad . \quad . \quad . \quad (12)$$

where each f_i is determined explicitly by expanding the canonical form F. For instance, a_1 and b_1 are two of these ν parameters l_i in the above binary cubic, and $f_1 = a_1{}^3 + b_1{}^3$.

Then if we can solve these N equations for N of the ν parameters in terms of the rest, the form F is legitimately canonical. If not, F is uncanonical. Now a solution is, or is not, possible according as a relation $\psi(f) = 0$ does not, or does, exist: that is, if the rank of the matrix $\left[\dfrac{\partial f_i}{\partial l_j}\right]$ is N, so that at least one N-rowed Jacobian $\dfrac{\partial(f)}{\partial(l)}$ does not vanish identically for all values of its N parameters l, then a canonical form exists.

But in the uncanonical case a relation $\psi(f) = 0$ must exist for all values of the ν parameters l. Thus

$$\frac{\partial f_1}{\partial l_r}\frac{\partial \psi}{\partial f_1} + \frac{\partial f_2}{\partial l_r}\frac{\partial \psi}{\partial f_2} + \ldots + \frac{\partial f_N}{\partial l_r}\frac{\partial \psi}{\partial f_N} = 0, \quad (r = 1, 2, \ldots, \nu).$$

Now this is a lineo-linear relation of type (8) between the N coefficients $\left[\dfrac{\partial f_1}{\partial l_r}, \ldots, \dfrac{\partial f_N}{\partial l_r}\right]$ of a form χ_r and the coefficients $\left.\dfrac{1}{c_1}\dfrac{\partial \psi}{\partial f_1}, \ldots, \dfrac{1}{c_N}\dfrac{\partial \psi}{\partial f_N}\right\}$ of a dual form $\phi(u)$. And owing to relations (12), this form χ_r is precisely $\dfrac{\partial f(x)}{\partial l_r}$.

Then unless the ν forms χ_r have an apolar form $\phi(u)$ for all values of the parameters, $\psi(f) = 0$ cannot exist and the form F is therefore canonical.

Finally, if any of the parameters are implicit, say $l = a_i$ in a linear form,

$$X = a_1 x_1 + a_2 x_2 + \ldots + a_n x_n,$$

then $$\partial f(x)/\partial l = \partial F(X)/\partial l = x_i \partial F/\partial X.$$

Hence for each value of i, $x_i \partial F/\partial X$ is apolar to $\phi(u)$. This requires $\partial F/\partial X$ to be apolar to $\phi(u)$. Conversely if $\partial F/\partial X$ is apolar, so is $\partial f(x)/\partial l$. Similarly if X is a form in $x_1, \ldots, x_n$ of higher order. This proves the theorem.

EXAMPLES

1. A general ternary quartic cannot be expressed as the sum of five fourth powers because a quartic exists apolar to the five special cubic forms $\frac{1}{4}\partial F/\partial X_i = X_i^3$.

Proof.—

Through five points a conic can be drawn. This conic counted twice has these five for double points. Hence by Ex. **6**, p. 265, a quartic a_x^4 apolar to five cubes exists, and dually a quartic u_a^4 apolar to five cubes X_i^3 exists.

2. A general binary form of order $2k - 1$ can be expressed as the sum of k linear forms, each raised to the same power $2k - 1$.

3. Any binary p-ic f, apolar to every p-ic φ, which has a linear factor X repeated λ times, must itself contain this factor X repeated $p - \lambda + 1$ times.

[Combine Ex. **3**, p. 263; (iii), p. 264; and §6. Treating X, and another linear form Y, as new independent variables, then the last λ coefficients of φ are zero: whence the *last* $p - \lambda + 1$ of f must also vanish.

CHAPTER XVIII

INVARIANT EQUATIONS AND GRAM'S THEOREM

1. Expression of a Gradient by Coefficients of Covariants.

Let $f(A, x)$ typify one or more ground forms whose typical coefficient A is symbolized by a product $a_i a_j a_k \dots$, and whose transformed coefficients are indicated by an accent. Then (§6, p. 202) a single-term product P' of coefficients A' is symbolized by a product of factors a_i', and therefore of $a_\xi, a_\eta, \dots, b_\xi, \dots$.

Now consider the n columns $\xi, \eta, \dots, \omega$, of the matrix M which transforms x to x', as a set of n cogredient *points*, then P' is at once the symbol of a concomitant for the ground forms and these n points, because it consists entirely of inner products such as a_ξ. Further, let P' be expanded by a Gordan-Capelli series as

$$\begin{aligned} P' &= a_\xi^{\,i} a_\eta^{\,j} \dots b_\xi^{\,k} \dots \\ &= \Sigma \Delta_0 P_0 + |M| \Sigma \Delta_1 P_1 + |M|^2 \Sigma \Delta_2 P_2 + \dots \quad (1) \end{aligned}$$

where each Σ denotes a concomitant, and Δ_i is an aggregate of polar operators $\left(\eta \dfrac{\partial}{\partial \xi}\right)$, &c., involving pairs from among $\xi, \eta, \dots, \omega$. Each P_i involves $(n-1)$ at most of these cogredient sets, and $|M|$ denotes $(\xi\eta \dots \omega)$. We provisionally call each P_i a *covariant.* For binary forms $(n=2)$ this has the usual meaning, because P_i has now only one set ξ.

Identity (1) is true for all values of $\xi_1, \dots, \omega_n$. Taking the unit matrix in particular, when $\xi_1 = \eta_2 = \dots = \omega_n = 1$ and all the rest vanish, we obtain $a_\xi = a_i' = a_i$; so that each P_i becomes a coefficient in a certain covariant P_i, $|M|$ becomes unity, and P' becomes the original product of actual coefficients A (cf. §1, p. 226).

Finally, if we select a number of terms P, isobaric and therefore homogeneous, in each ξ, η, &c., and add the results, we obtain the theorem:

Any gradient can be expressed as the sum of coefficients in covariants involving n — 1, *or fewer cogredient variables.*

Examples.—

1. For binary cubic forms $(a_0, a_1, a_2, a_3 \between x_1, x_2)^3$, $(b_0, b_1, b_2, b_3 \between x_1, x_2)^3$ let $P = a_0 b_1$. Then

$$2P' = \tfrac{1}{3}\left(\eta\frac{\partial}{\partial\xi}\right) a_\xi^3 b_\xi^3 + (\xi\eta)(ab)a_\xi^2 b_\xi^2.$$

2. Let $\varphi(a)$ be a polynomial of binary coefficients such that when $\varphi(a) = 0$ so also $\varphi(a') = 0$ for all values of $\xi_1, \xi_2, \eta_1, \eta_2$. By taking the diagonal matrix $(\xi_2 = \eta_1 = 0)$ show that $\varphi(a)$ must be *isobaric.*

3. The Gordan-Capelli series for $\varphi(a')$ is

$$\varphi(a') = \Delta_0 P_0 + (\xi\eta)\Delta_1 P_1 + (\xi\eta)^2\Delta_2 P_2 + \dots$$

where the typical co-factor of $(\xi\eta)^r$ is a polar of a covariant, say c_ξ^p. Except for a numerical factor the typical term is $(\xi\eta)^r c_\xi^{p-m+r} c_\eta^{m-r}$. Putting $\xi_1 = \eta_2 = 1$, $\xi_2 = \eta_1 = 0$, then $\varphi(a') = \varphi(a)$ and the typical term is $c_1^{p-m+r} c_2^{m-r}$, which is a coefficient in the covariant c_ξ^p.

Hence *if* $\varphi(\mathrm{a}) = 0$ *and also* $\varphi(\mathrm{a}') = 0$ *for all transformations, a certain set of covariants* $P_0, P_1, \dots$ *must vanish identically.*

2. Invariant Equations.

Suppose a relation $\phi(A) = 0$ to exist between the coefficients of one or more general ground forms, in such wise that exactly the same relation $\phi(A') = 0$ exists for the corresponding coefficients after an arbitrary linear transformation. Then this relation is called an *invariant equation.* Geometrically, such relations are called *projective relations.*

It is clear that this is the kind of result which frequently occurs in analytical geometry (cf. p. 132); but it is by no means obvious that the case of such equations is covered by our invariant theory, because the condition $I(A') = |M|^w I(A)$ of the latter is more stringent than to say $\phi(A) = 0 = \phi(A')$. Nevertheless they are closely related, as the following theorem demonstrates.

3. Gram's Theorem.[1]

If an invariant equation exists it belongs to a system of m *such equations which specify that the* m *coefficients of a certain covariant, in* n *or less variables cogredient with* x, *vanish. If* m = 1 *the equation specifies the vanishing of an invariant.*

---

[1] *Math. Annalen*, **7** (1873), 230–240.

Conversely, if a covariant vanishes identically it continues to do so after linear transformation.

Proof.—

This converse is obvious, while the direct theorem follows from the result of the last section.

For if $\phi(A) = 0$ is an invariant equation then by definition $\phi(A')$ vanishes identically for all values of the transformation coefficients $\xi_1, \ldots, \omega_n$. We construct $\phi(A')$ as a function of the A's and the ξ's, η's, . . . , expressing it *in its lowest terms* as

$$\phi(A') = c_1\phi_1(A) + c_2\phi_2(A) + \ldots + c_m\phi_m(A), \qquad (2)$$

where each c_i is a function of $\xi, \eta, \ldots$ only. As $\phi(A')$ vanishes for all values of $\xi_1, \ldots, \omega_n$ it follows that

$$\phi_1(A) = 0, \quad \phi_2(A) = 0, \quad \ldots, \quad \phi_m(A) = 0. \qquad (3)$$

Thus we deduce m linearly independent conditions as a consequence of an invariant equation $\phi(A) = 0$. Interchanging the rôles of A and A' we deduce from the inverse transformation,

$$\phi_1(A') = 0, \quad \phi_2(A') = 0, \quad \ldots, \quad \phi_m(A') = 0. \qquad (4)$$

Thus we have a system of m invariant equations. But so far we have not shown that they include $\phi(A)$. This follows by substituting $\xi_1 = \eta_2 = \ldots = \omega_n = 1$, &c., from the identical transformation, making $\phi(A') = \phi(A)$. For now (2) becomes

$$\phi(A) = \lambda_1\phi_1(A) + \lambda_2\phi_2(A) + \ldots + \lambda_m\phi_m(A), \qquad (5)$$

where each λ is numerical.

Further, by making the transformation matrix M a *diagonal* matrix (Ex. **1**, p. 101) with zeros everywhere but on the diagonal $\xi_1, \eta_2, \ldots, \omega_n$ and substituting these values of $\xi, \ldots, \omega$ in (2) we infer that parts of $\phi(A')$ homogeneous separately in each $\xi, \eta, \ldots$ satisfy the typical invariant equation condition. Hence we lose no generality by assuming $\phi(A)$ is homogeneous and isobaric in A. In fact $\phi(A)$ is a gradient.

Finally by the Gordan-Capelli expansion for $\phi(A')$ of the preceding section we have an explicit form for condition (2), which at once shows that the m functions $\phi_i(A)$ are the several distinct coefficients in a covariant, or combination of covariants.

In particular if $m = 1$, then, by (5), $\phi(A) = \lambda_1 \phi_1(A)$, and by (2), $\phi(A') = \frac{c_1}{\lambda_1} \phi(A)$, showing at once that $\phi(A)$ is an invariant. This proves the theorem.

4. Grace's Theorem.

Quite recently Mr. J. H. Grace[1] has developed this theory with more particular reference to *binary* vanishing covariants. The question which is put now runs as follows:

What is the most general polynomial $\phi(\mathrm{A})$ *of degree* i *in the coefficients* A *which vanishes when the ground forms have an assigned projective property?*

The answer is simple, namely:

$\phi(\mathrm{A})$ *is the sum of a number of parts each of which is a coefficient in a covariant of the forms, and all such covariants vanish in virtue of the assigned property.*

Such covariants have already been found; they can be taken as linearly independent of degree i.[2] But it can further be proved that the coefficients themselves are linearly independent. In fact, there cannot be a linear relation between any coefficients of a set of linearly independent covariants, for if there were, the operation of writing a_ξ for a_1, a_η for a_2, &c., would immediately give a linear relation between the covariants themselves. But the argument breaks down for ternary and higher forms.

This theory is applicable to forms of all types already contemplated, including multiple fields. Little or nothing is known about these last, and in fact the whole subject presents many opportunities for further investigation.

Examples.—

1. *To find the necessary and sufficient conditions for a ternary cubic to be a perfect cube.*

Let $a_x^3 = b_x^3$ be the symbolic form of the cubic which is to be a perfect cube $(p_1x_1 + p_2x_2 + p_3x_3)^3$ where the coefficients p_i are actual numbers, and a_{ijk} is the actual coefficient in non-symbolic form. Thus we have a number of non-symbolic equations

$$a_{ijk} = p_i p_j p_k \qquad (i, j, k = 1, 2, 3).$$

This is secured by eliminating p_1, p_2, p_3 in every possible way, giving

$$a_{ijk}^2 = a_{iij} a_{jkk}, \qquad a_{iii} a_{jjj} = a_{iij} a_{ijj}.$$

---

[1] *Journal London Math. Soc.*, **3** (1928), 34–38, **5** (1929), 62–67.

[2] This remark as to the linear independence of the coefficients applies only to *binary* forms. Cf. J. H. Grace, *Journal London Math. Soc.*, **5** (1930), 65.

Symbolically this gives two types of condition

$$a_1 a_2 a_3 \,.\, b_1 b_2 b_3 = a_1^2 a_2 \,.\, b_2 b_3^2; \quad a_1^3 \,.\, b_2^3 = a_1^2 a_2 \,.\, b_1 b_2^2.$$

The first leads by Gram's theorem to the covariant equation

$$a_x a_y a_z b_x b_y b_z - a_x^2 a_y b_y b_z^2 = a_x a_y b_y b_z [a_z b_x - a_x b_z] = 0.$$

If $u_i = (zx)_{jk}$ we can write this

$$(abu)\, a_x a_y b_y b_z,$$

and by interchange of equivalent symbols this becomes

$$\tfrac{1}{2}(abu)(ab \mid xz)\, a_y b_y = \tfrac{1}{2}(abu)(abv)\, a_y b_y,$$

where v is cogredient to u. This concomitant is a polar of

$$\tfrac{1}{2}(abu)^2 a_x b_x \equiv \tfrac{1}{2}\Theta.$$

The other condition gives the covariant

$$\begin{aligned} a_x^3 b_y^3 - a_x^2 a_y b_x b_y^2 &= a_x^2 b_y^2 (ab \mid xy) \\ &= \tfrac{1}{2}(abu)(a_x^2 b_y^2 - a_y^2 b_x^2) = \tfrac{1}{2}(abu)^2 (a_x b_y + a_y b_x), \end{aligned}$$

which is a polar with regard to y of the same form Θ.

Hence the required condition for a_x^3 to be a perfect cube is that the mixed concomitant Θ should vanish identically.

2. The cubic a_x^3 in n variables is a perfect cube if $(ab \mid xy)^2 a_x b_x$ vanish identically.

3. The quadratic a_x^2 in n variables is a perfect square if $(ab \mid xy)^2$ vanish identically.

4. The binary n-ic is a perfect nth power if the Hessian vanish identically: and consequently all its concomitants except itself vanish.

5. For the binary n-ic which contains a factor repeated $n - 1$ times all covariants of grade four, i.e. such as contain the symbolic factor $(ab)^4$, vanish. (*Grace.*)

6. Let $f = a_x^n = b_x^n$ be a binary n-ic. Show that a complete set of covariants of degree two is

$$f^2, \quad H_1 = (ab)^2 a_x^{n-2} b_x^{n-2}, \quad H_2 = (ab)^4 a_x^{n-4} b_x^{n-4},$$
$$H_3 = (ab)^6 a_x^{n-6} b_x^{n-6}, \quad \&c.$$

7. If l_x is an actual linear factor, repeated $n - i$ times in f, so that

$$f \equiv 0 \bmod l_x^{n-i},$$

then all covariants of degree two vanish except

$$f^2, \quad H_1, \quad H_2, \ldots, H_i.$$

(*Grace.*)

5.[1] Invariants as Elimination Results.

Let $f = a_x^p$ be an n-ary p-ic whose actual coefficients are now written

$$a_1, a_2, \ldots, a_N. \quad \ldots \ldots \ldots \quad (6)$$

Further, let the linear transformation $x \longrightarrow x'$, of matrix $[e_{ij}]$, induce a transformation $a \rightarrow b$ on these coefficients, so that the new coefficients are given by N linear equations

---

[1] Gram, *Math. Annalen, loc. cit.*

$$b_h = \sum_{\rho=1}^{N} a_\rho \phi_{\rho h} (e_{ij}). \tag{7}$$

Then, if f has two polynomial invariants $I(a)$, $K(a)$ of the same weight, they give rise to an absolute invariant

$$i(a) = \frac{I(a)}{K(a)}. \tag{8}$$

Thus

$$\frac{I(a)}{K(a)} - \frac{I(b)}{K(b)} = 0, \quad \text{or} \quad I(a)\,K(b) - K(a)\,I(b) = 0. \tag{9}$$

This last is a polynomial equation in the coefficients a_i, b_h which by definition must vanish identically for all values of e_{ij} when each b_h is expressed in terms of the original a's. In other words (9) is the result of eliminating the n^2 coefficients e_{ij} from the N equations (7). *Each absolute invariant is an elimination result, or let us simply say a resultant, of the system* (7) *regarding* e_{ij} *as the variables.*

For this to happen in general the number N of equations (7) must exceed n^2, else the e's cannot be eliminated. So we assume $N > n^2$. But we can go further and prove conversely that

Every resultant derived by eliminating the n^2 *coefficients* e_{ij} *from the* N *transformation equations furnishes, either a system of invariant equations for the coefficients* a_h *of the ground form* f, *or a relation* $i(a) = i(b)$ *expressing the equality of two absolute invariants.*

Proof.—

Let such a resultant be expressed as a polynomial in each a_h and b_h as

$$R(a, b) = 0. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \tag{10}$$

There are two cases to consider. Either one set of coefficients is absent from R, or both are present.

First we can take R to be $R(a)$, excluding b. By interchanging a, b and using the inverse transformation and carrying out identically the same steps we should have arrived at the analogous result $R(b) = 0$. Hence $R(a) = 0$ belongs to an invariant system of equations, and by Gram's theorem certain invariants or covariants of f vanish identically. In this case f is not a general n-ary p-ic.

Secondly, R contains coefficients of both types a and b. We write the resultant explicitly as

$$R \equiv A_0'B_0 + A_1'B_1 + \ldots + A_\nu'B_\nu = 0, \quad . \quad (11)$$

where, for each value of i, A_i' is a homogeneous polynomial of degree i in the coefficients a, and B_i likewise in b. Further we suppose each such resultant to be in its simplest terms and irresoluble into factors.

Let a new linear transformation $x \rightarrow y$ induce the coefficient transformation $a \rightarrow c$ of matrix $[\epsilon_{ij}]$. Then between sets (a) and (c) there will be a corresponding resultant

$$R' \equiv A_0'C_0 + \ldots + A_\nu'C_\nu = 0, \quad . \quad . \quad (12)$$

each C_i being the same function of (c) that B_i is of (b). And since (a), (b), (c) are connected linearly, there will be a matrix $[\theta_{ij}]$ for the direct transformation $b \rightarrow c$, with its corresponding resultant

$$R'' \equiv B_0'C_0 + \ldots + B_\nu'C_\nu = 0, \quad . \quad . \quad (13)$$

where each B_i' is analogous to the original A_i'.

Solving (11) and (12) for A_0' and equating the results, we have

$$A_1'\frac{B_1}{B_0} + \ldots + A_\nu'\frac{B_\nu}{B_0} = A_1'\frac{C_1}{C_0} + \ldots + A_\nu'\frac{C_\nu}{C_0}, \quad (14)$$

which along with each preceding relation must be an identity for all values of the elements e_{ij}, ϵ_{ij}, θ_{ij}. Then if we express each C_i as a function of (b) and (θ), the result (14) gives a relation between sets (a), (b) also involving an *arbitrary* set (θ). As the a's only enter (14) by way of the ν homogeneous polynomials A_i', and, by (7), the b's are arbitrary, it follows that the coefficients of A_i' are equal on each side. Thus

$$\frac{B_1}{B_0} = \frac{C_1}{C_0} \qquad \frac{B_2}{B_0} = \frac{C_2}{C_0}, \quad \&c.$$

Hence by interchanging the rôles of (a) and (c) we have $\frac{A_i}{A_0} = \frac{B_i}{B_0}$, $i = 1, 2, \ldots, \nu$, each of which is a relation of type

$$\frac{\phi(a)}{\psi(a)} = \frac{\phi(b)}{\psi(b)}.$$

In other words it is an equality between two absolute invariants $i(a) = i(b)$.

Corollary I.—*The numerator and the denominator of an absolute invariant* $\phi(\mathrm{a}) / \psi(\mathrm{a})$ *are relative invariants.*

For if we write the absolute relation as

$$\phi(b) = \rho\phi(a), \quad \psi(b) = \rho\psi(a),$$

then $\phi(b)$ and $\psi(b)$ are polynomials in both sets a and e_{ij}. Consequently ρ is a rational function

$$\frac{p(a, e)}{q(a, e)}$$

certainly involving the set e and possibly a. If we suppose p/q to be in its lowest terms, then, in order to make $\rho\phi(a)$ a polynomial, $q(a, e)$ must be a factor of $\phi(a)$, and similarly of $\psi(a)$. This is impossible, if $\phi(a) / \psi(a)$ is originally in its lowest terms, unless $q(a, e)$ is a mere constant factor. Hence ρ is a polynomial.

Further, since each b is linear in the set a, $\phi(b)$ and $\phi(a)$ are of the same degree in a. Hence the degree of ρ is zero, so that ρ depends solely on the set e_{ij}. By the definition (§2, p. 169) this proves the result.

Corollary II.—*The theorem holds for any simultaneous system of ground forms.*

6. The Equivalence Problem.

When a linear transformation $x \longrightarrow x'$, with non-vanishing determinant $| e_{ij} | = | M |$ turns a form f into f' and consequently the inverse transformation $x' \longrightarrow x$ turns f' into f, the two forms are said to be equivalent. Manifestly if f is equivalent to f' and f' to f'' then f is also equivalent to f''.

When the transformation changes f to $\rho f'$ where ρ is a non-zero constant, the *equations* $f = 0, f' = 0$ are said to be equivalent. It is easy to adapt this last to the original case by multiplying each e_{ij} by the same constant $\sqrt[m]{\rho}$.

The results of the foregoing sections show the necessary and sufficient conditions for two forms to be equivalent. The coefficients (a) and (b) of the equivalent forms must either both be general, or else both satisfy the same particular conditions, and also their corresponding absolute invariants must be equal. By Gram's theorem this first entails the identical vanishing of the same covariants.

7. Extension of Stroh's Lemma.

Recently Mr. J. H. Grace[1] has used the preceding methods of canonical forms, with great effect, for theorems which used to be very difficult to prove, though of considerable importance in applying the fundamental identities

$$(bc)a_x + (ca)b_x + (ab)c_x = 0$$
$$(bcd)a_x + (cad)b_x + (abd)c_x + (bac)d_x = 0$$

to binary and ternary forms.

For *if* $\lambda_1, \lambda_2, \ldots, \lambda_r$ *are positive integers, each not exceeding* p, *such that*

$$\lambda_1 + \lambda_2 + \ldots + \lambda_r = (r-1)(p+1), \quad r > 2, \quad (1)$$

then a canonical form of f(ξ), *any general or special binary* p-*ic in* ξ_1, ξ_2, *is*

$$f(\xi) = X_1^{\lambda_1}P_1 + X_2^{\lambda_2}P_2 + \ldots + X_r^{\lambda_r}P_r, \quad . \quad (2)$$

where $X_1, X_2, \ldots X_r$ *are* r *given distinct linear forms in* ξ_1, ξ_2, *and* P_i *is a form of order* p $-\lambda_i$. *For a given set of* X's *this set of* P's *is unique.*

Proof.—

Regarding the coefficients in P_i as intrinsic parameters of the proposed canonical form, we note that P_i supplies μ_i such parameters where $\mu_i = p - \lambda_i + 1$. So the P's supply in all,

$$\sum_1^r (p - \lambda_i + 1) = r(p+1) - \Sigma\lambda_i = p + 1, \quad . \quad (3)$$

a number which tallies with the number of coefficients in $f(\xi)$.

Again, since (2), when written in full, is linear in these parameters, not only does it provide $p+1$ equations for the $p+1$ parameters, but the equations also are linear. Hence the canonical form is *unique*, if it exist.

Now suppose $\phi(\xi)$ is a binary p-ic always apolar to $\sum_{i=1}^r X_i^{\lambda_i}P_i$.

---

[1] *Proc. Cambridge Phil. Soc.*, **24** (1928), 218–222.

Then, if all $p-\lambda_i+1$ coefficients in each P_i are arbitrary, $\phi(\xi)$ will be apolar to $X_i^{\lambda_i}P_i$ singly, and therefore (Ex. 3, p. 269) will have $X_i^{\mu_i}$ as factor. Hence $\phi(\xi)$ will contain distinct factors $X_1^{\mu_1}, X_2^{\mu_2}, \ldots X_r^{\mu_r}$ giving by (3) a total order *greater* than p, its own order: which is impossible. Consequently no such apolar form $\phi(\xi)$ exists, and the forms P_i are entirely linearly independent, so that the canonical form (2) is justified.

Corollary I.—If $r=3$, we obtain *Stroh's lemma:*

If ξ, η, ζ are three quantities whose sum is zero, and λ, μ, ν three positive integers ($\leqslant$ p) *such that*

$$\lambda+\mu+\nu=2p+2$$

then any homogeneous polynomial of order p *in ξ, η, ζ can be expressed uniquely in the form*

$$\xi^\lambda P+\eta^\mu Q+\zeta^\nu R.$$

For let $X_1=\xi$, $X_2=\eta$, $X_3=-\xi-\eta$, and $r=3$ in the above theorem.

Corollary II.—Further, if as is possible, $\lambda \geqslant \frac{2}{3}p$, $\mu \geqslant \frac{2}{3}p$, $\nu \geqslant \frac{2}{3}p$ we obtain what is known as *Jordan's lemma.*

Similar methods apply[1] to four quantities ξ, η, ζ, ω, whose sum is zero. Any p-ic in ξ, η, ζ, ω can be expressed (not necessarily uniquely) as

$$\xi^\lambda P+\eta^\mu Q+\zeta^\nu R+\omega^\rho S$$

where $\lambda+\mu+\nu+\rho=2p+3$. For five variables $2p+3$ changes into $2p+4$, and so on.

EXAMPLE

Prove the identity

$$3(bc)(ca)(ab)a_x^2b_x^2c_x^2=(bc)^3a_x^3+(ca)^3b_x^3+(ab)^3c_x^3.$$

---

[1] *Loc. cit.*

CHAPTER XIX

GEOMETRICAL INTERPRETATIONS OF ALGEBRAIC FORMS

1. Homogeneity and Correspondence.

In Chapters I and V allusions have been made to Cartesian and homogeneous co-ordinates. We now seek a closer connexion between the geometry and the algebra. The straight line, with its totality of points, illustrates the binary theory—a finite set of n points picturing the binary n-ic—while the points of a plane illustrate ternary forms, the points of threefold space, quaternary forms, and so on. Moreover it is worth while pondering for a moment on two relevant ideas which help to make the general setting of the theory a little clearer. One is the idea of *homogeneity*, and the other is that of *correspondence.*

(1) The practice of centuries in algebra has made it abundantly clear that the homogeneous polynomial, or form, is much easier to handle than the non-homogeneous. And although in ordinary Cartesian geometry a start is usually made with the non-homogeneous expressions, as in the general equation of the second degree for a conic, in the natural, but inflated, hope that such a course is more comprehensive and general, very soon we return to the homogeneous once more, either by paying attention to the terms of highest degree

$$ax^2 + 2hxy + by^2,$$

or by introducing a third variable z, and considering what is in effect the *ternary quadratic form*

$$\begin{aligned} S &= ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy \\ &\equiv (a, b, c, f, g, h \between x, y, z)^2. \end{aligned}$$

A return to Cartesian co-ordinates can then be made at any moment merely by putting $z = 1$.

Likewise the general polynomial F of order p in $n-1$ variables can be written

$$F = U_0 + U_1 + \ldots + U_n,$$

where each term U is a homogeneous form in the variables, of order indicated by the suffix. So if t is a new variable, this form is made homogeneous in n variables by writing

$$U_0 t^p + U_1 t^{p-1} + \ldots + U_p,$$

taking F as the special case of this, when $t = 1$. So in discussing the n-ary form—the polynomial homogeneous in n variables—we are including the apparently more general non-homogeneous form.

It is only when $U_0 + U_1 + \ldots + U_p$ is an endless series that homogeneity breaks down; and it is just here that we step over the clear border line between algebra and analysis. Algebra is, in fact, the study of the finite, wherein it is totally different from elementary arithmetic on the one hand and analysis on the other. Thus at once we begin to express ourselves algebraically when we write x, a single finite symbol for the endless choice of positive integers which form the ambitious subject matter of elementary arithmetic.

It should, however, be remarked that, in analysis, homogeneity can be retained; but only at the expense of another algebraic feature—the polynomial. For example, the series

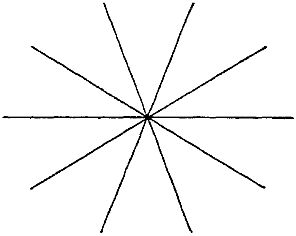
$$1 + \frac{x}{y} + \ldots + \frac{x^p}{y^p} + \ldots$$

is homogeneous in two variables x, y; and each term is rational, but not integral.

(2) *Correspondence.*—This is perhaps only another name for the same idea. Just as x is one symbol for a numerous class of things, so an algebraic matrix, form, equation, or syzygy is one symbol for many phenomena in quite divergent fields of thought. A very simple familiar example of this is given by the two geometrical figures on the following page, one with points in line, the other with coplanar lines through a common point. As objects to be gazed upon what could be more different?

Yet they contain the same idea; and algebraically they are implied by the same symbols.

In short it may be said that there is not merely one but a definite system of geometrical and even physical phenomena associated with each algebraic statement, and no hard and fast rule holds for the geometrical interpretation of the algebra.



2. Principle of Duality.

The above two figures, of points in line, and lines through a point, illustrate in geometry what is called the principle of duality or reciprocation. This principle immediately fits in with the duality already seen in the algebra—in the rows and columns of a matrix, in the theory of reciprocal matrices and determinants, and briefly in all that is comprised in the terms cogredience and contragredience. It amounts to this: that just as the algebra attained greater richness and completeness by the use of two sorts of variables x and u, contragredient to one another in a field of order n, so also geometry in any number of dimensions (say $n - 1$ dimensions as equivalent to a field of order n), becomes more intelligible by the use of two sorts of elements—points and primes. A prime is a space of $(n - 2)$ dimensions relative to the field of order n. Prime is a useful word because it covers various cases: thus a prime is a line in a plane field (when $n = 3$); it is a plane in threefold space (when $n = 4$); and so on. Then it is found that every geometrical property of points in the field can be matched by a corresponding property of primes; and such are called dual or reciprocal properties. At present it is enough to notice that there are two kinds of reciprocal properties, the one arising from the very fundamental texture of space and the second arising from a ground form Γ, now interpreted as a geometrical locus, such as a conic, for which pole and polar elements exist.

To illustrate these remarks consider the ternary case. A triangle ABC may be thought of as a set of three points A, B, C or a set of three lines BC, CA, AB. In this we have an example of the first kind of duality. As further instances of such properties we can set, side by side, the facts:

Two sides a, b of the triangle ABC pass through one point C.	Two points A, B of the triangle abc lie on one line c.

But we obtain similar dual results by taking a conic Γ in the plane of a triangle ABC and forming the polars of the points A, B, C with regard to the conic. This usually gives a new triangle, say $a'b'c'$, where a' is polar of A, b' of B, and c' of C. It is now quite easy to write down dual properties, the one holding for the triangle ABC, and the other for $a'b'c'$.

EXAMPLES

1. The binary form $(a_0, a_1, \ldots, a_p \between x_1, x_2)^p$ equated to zero, represents p points on a line, points which may be real, coincident or complex. Each root for $x_1 : x_2$ gives one such point P, and $x_1 : x_2$ may be interpreted as a ratio determining the position of P relative to two fixed base points B_1, B_2 of the line, in the familiar elementary way.

2. Non-homogeneously, if $x_1 = x$, $x_2 = 1$, we interpret this binary p-ic by the use of a Cartesian co-ordinate x, relative to a given origin O on the line.

3. Four binary linear forms a_x, b_x, c_x, d_x have an absolute invariant

$$k = \frac{(ab)(cd)}{(ad)(cb)} = \{abcd\}.$$

This is called the *cross ratio* or *anharmonic ratio* of the four forms. By interchanging a, b, c, d in all 24 ways, derive six cross ratios.

Ans. $k,\ 1 - k,\ \frac{1}{k},\ \frac{1}{1-k},\ 1 - \frac{1}{k},\ \frac{k}{k-1}$

These follow by using the identity $(bc)(ad) + (ca)(bd) + (ab)(cd) = 0$.

4. Prove $\{abcd\} = \{cdab\} = \{dcba\} = \{badc\}$.

5. Prove that the operations of deriving the remaining five from any one of the above six cross ratios form a group.

6. Prove that $\{abcd\}$ denotes the geometrical cross ratio of four points in a line. Here $a_x = 0$ gives the point $(a_2, -a_1)$, in homogeneous co-ordinates.

7. Examine the special cases when (i) $k = 0, 1, \infty$; (ii) $k = -1, 2$ or $\frac{1}{2}$.

[(i) Two points coincide. (ii) The range is harmonic.

8. If the roots of the quadratics $a_x^2 = 0$, $b_x^2 = 0$ denote two pairs of points, prove that they separate each other harmonically if $(ab)^2 = 0$.

Non-symbolically if $a_0 b_2 + a_2 b_0 - 2a_1 b_1 = 0$, then $a_0 x^2 + 2a_1 x + a_2$ and $b_0 x^2 + 2b_1 x + b_2$ determine harmonic pairs of points.

9. *The Jacobian* $(\mathrm{ab})\mathrm{a_x b_x} = 0$ *determines two points* K, L *which are a harmonic pair, simultaneously for the pairs* P, Q *and* R, S, *given respectively by* $\mathrm{a_x}^2 = 0$, $\mathrm{b_x}^2 = 0$.

10. If $(bc)\,(ca)\,(ab) = 0$ (§4, p. 218) the quadratics a_x^2, b_x^2, c_x^2 are each harmonic to a common quadratic j_x^2.

11. All quadratics of a pencil $\lambda f + \lambda' f'$ have a common harmonic quadratic.

Ans. The Jacobian (f, f').

3. Further Binary Results.[1]

A second interpretation of binary forms dual to Example 1 above is to treat the variables x_1, x_2 as ordinary Cartesian co-ordinates and the binary n-ic as representing n straight lines through the origin. In this case two quadratics represent two pairs of lines through a fixed point, and the vanishing of their simultaneous invariant now gives the necessary and sufficient condition for these pairs to form a harmonic pencil.

EXAMPLE

Prove that such pairs of lines meet any arbitrary line, which does not contain the fixed point O, in a harmonic range.

A third interpretation of binary forms is to treat the n-ic as representative of n points on a rational plane curve, by taking the variable $x_1 : x_2$ as the parameter of a point of the curve. For example, if (X, Y) are Cartesian co-ordinates, then the relations

$$X : Y : 1 = x_1^2 : x_1 x_2 : x_2^2$$

determine the conic $X = Y^2$, and a binary n-ic in x_1, x_2, equated to zero, gives n points on the conic. If ternary homogeneous co-ordinates X, Y, Z are used, the conic $XZ = Y^2$ has the parametric equations

$$X : Y : Z = x_1^2 : x_1 x_2 : x_2^2. \quad . \quad . \quad . \quad . \quad (1)$$

It can be proved that projective properties of points on the conic correspond to binary invariants and covariants.

---

[1] For further treatment, the reader should consult Grace and Young, *Algebra of Invariants*, Chap. X.

A fourth interpretation of binary forms is the dual of the third. The n-ic represents n tangents to a rational curve. Thus if $UX + VY = 1$ is the equation of a line in Cartesian co-ordinates, U, V are called *line-* or *tangential* co-ordinates. Then if

$$U : V : 1 = x_1^2 : x_1x_2 : x_2^2,$$

the line touches a conic whose tangential equation is $U = V^2$, and whose point equation is $Y^2 + 4X = 0$. Similarly for homogeneous line co-ordinates.

EXAMPLES

1. If $X : Y : Z$
$= ax_1^2 + 2hx_1x_2 + bx_2^2 : a'x_1^2 + 2h'x_1x_2 + b'x_2^2 : a''x_1^2 + 2h''x_1x_2 + b''x_2^2$,
show that a ternary linear transformation $X, Y, Z \longrightarrow X', Y', Z'$ in general exists such that $X' : Y' : Z' = x_1^2 : x_1x_2 : x_2^2$.

Show that both points (X, Y, Z) and (X', Y', Z') lie on conics, for all values of x_1, x_2.

2. If $X : Y : Z = a_x^n : b_x^n : c_x^n$, then (X, Y, Z) lies on a plane curve of order n, which is rational.

[The line $UX + VY + WZ = 0$ cuts the curve in n points. Rational, because this parametric form is rational.]

A fifth interpretation is to consider the n roots of a binary n-ic to be points in the Gauss plane. This method has the advantage of giving a real geometrical figure for complex binary forms.

Lastly a sixth interpretation, and probably the most profound, is by means of the norm curve in space of n dimensions of which the plane conic (1) illustrates the quadratic case. By this is meant; taking

$$X_1 : X_2 : \ldots : X_{n+1} = x_1^n : x_1^{n-1}, \qquad : x_2^n \qquad (2)$$

where $(n + 1)$ co-ordinates X_1, X_2 $\qquad$ X_{n+1} are called the homogeneous co-ordinates of a point in n-fold space. We gain a hint of its possibilities by noticing that, if $n = 3$, a point (X_1, X_2, X_3, X_4) of ordinary three-fold space lies on a curve which meets an arbitrary plane

$$A_1X_1 + A_2X_2 + A_3X_3 + A_4X_4 = 0$$

in three points, if its co-ordinates satisfy (2). For the ratio

$x_1 : x_2$ then can only take three values, which are the roots of a binary cubic

$$A_1 x_1^3 + A_2 x_1^2 x_2 + A_3 x_1 x_2^2 + A_4 x_2^3 = 0.$$

Such a curve is called a twisted cubic. If we think of this norm curve as fixed in space, then each binary cubic is associated with a plane in space.

For example, if a binary cubic has a repeated factor its plane touches the cubic curve; if it is a perfect cube, its plane osculates the curve.

4. Connexion of Binary with Higher Fields.

If, for ternary forms, y and z are two distinct points $\{y_1, y_2, y_3\}$, $\{z_1, z_2, z_3\}$, then $\{\xi_1 y_1 + \xi_2 z_1, \xi_1 y_2 + \xi_2 z_2, \xi_1 y_3 + \xi_2 z_3\}$ are the co-ordinates of any point X in the line yz. Further, if $u_x = 0$ is the equation of an arbitrary line, then the point X lies on it if

$$u_1(\xi_1 y_1 + \xi_2 z_1) + u_2(\xi_1 y_2 + \xi_2 z_2) + u_3(\xi_1 y_3 + \xi_2 z_3) = 0, \quad (3)$$

which can be written shortly as

$$u_X \equiv u_y \xi_1 + u_z \xi_2 = 0. \quad (4)$$

Similarly if $a_x^p = 0$ denotes a curve of order p, the line yz cuts it at points X, for which ξ_1, ξ_2 are given by the binary p-ic

$$\begin{aligned} a_X^p &\equiv (a_y \xi_1 + a_z \xi_2)^p = 0, \\ \text{i.e.} \quad a_y^p \xi_1^p &+ p\, a_y^{p-1} a_z \xi_1^{p-1} \xi_2 + \ldots + a_z^p \xi_2^p = 0. \end{aligned} \quad (5)$$

It is essential to notice that formulæ (4) and (5) are precisely the same if we start with two points y, z in space of *any* dimension; for the extra terms in (3) are implied by the inner products u_y, a_y, &c., of (4) and (5).

The theory of tangents and polars of conics, quadrics, and loci of higher orders can be readily deduced from (5) by the ordinary elementary methods.

EXAMPLES

1. The equation of the tangent to a conic $a_x^2 = 0$ at the point y on the curve is $a_x a_y = 0$.

[Put $p = 2$, $a_y^2 = 0$ in (5), and make the two roots for $\xi_1 : \xi_2$ equal.

2. The polar of y is $a_x a_y = 0$.

3. The tangent prime at the point y to the quadric $a_x^2 = 0$ in general is $a_x a_y = 0$. This equation also gives the polar of y.

4. If $a_y a_z = 0$ the points y, z are conjugate. Prove that the line yz cuts the conic (or quadric) in two points harmonically separating y, z.

5. The line yz touches $a_x^2 = 0$ if $a_y^2 b_z^2 - a_y a_z b_y b_z$ vanishes, this being the discriminant for equal roots $\xi_1 : \xi_2$. The symbols a, b are equivalent.

This reduces to $(ab \mid yz)^2 = 0$.

6. In ternary forms if $u = yz$, then $u_x = (xyz)$; and $u_x = 0$ or $(xyz) = 0$ is the equation of the line yz. Prove that this line u touches the conic $f = a_x^2 = b_x^2 = 0$, if $(abu)^2 = 0$.

7. If $f = \sum\limits_{i,j=1}^{3} a_{ij} x_i x_j$, $a_{ij} = a_{ji}$, then

$$\left.\begin{array}{cccc} \cdot & u_1 & u_2 & u_3 \\ u_1 & a_{11} & a_{12} & a_{13} \\ u_2 & a_{21} & a_{22} & a_{23} \\ u_3 & a_{31} & a_{32} & a_{33} \end{array}\right| \equiv -\tfrac{1}{2}(abu)^2.$$

8. Write out the dual statement of this §4 and these examples. In particular, if two lines u, v are conjugate for the conic $(abu)^2 \equiv u_\alpha^2 = : 0$, then $u_\alpha v_\alpha = 0$.

5. The Clebsch Transference Principle. Extensionals.

We have met with the elegant theory of extensionals (Corollary III, p. 49) wherein a general property of n-rowed determinants leads to analogous properties of determinants of higher order. This conception influences the symbolic invariant theory, particularly *in exhibiting the actual working of projection from one space to another*. As a rule the methods of algebraic and pure geometry are alien to each other, only having something in common at the beginning and end of a chain of reasoning. But here they are in close touch, and furnish one of the beautiful harmonies of mathematics. The principle is due to Clebsch.

An illustration leading to this principle is given by the work of the preceding section. In fact we can look on the expression $a_x = a_y \xi_1 + a_z \xi_2$, used in (5), as a symbolic *binary* form in variables ξ_1, ξ_2, if we treat a_y, a_z as binary symbols A_1, A_2 by taking.

$$A_1 = a_y, \quad A_2 = a_z, \quad \text{so that } A_\xi = a_x. \qquad (6)$$

Thus in Ex. **5** above, the original quadric in X is now A_ξ^2; and if equivalent symbols B are used, the condition for the line yz to touch the quadric now reads as

$$\tfrac{1}{2}(AB)^2 = \tfrac{1}{2}(a_y b_z - a_z b_y)^2 = \tfrac{1}{2}(ab \mid yz)^2 = 0. \quad . \quad (7)$$

This shows that from a certain binary invariant (usually symbolized by $(ab)^2$ and here by $(AB)^2$) we deduce a concomitant involving an arbitrary line yz in higher dimensions, such that the points common to a quadric and the line coincide if the binary invariant vanishes.

By the principle of duality we can also write

$$\tfrac{1}{2}(AB)^2 = \tfrac{1}{2}(abuv\ldots w)^2$$

in terms of $(n-2)$ primes $u, v, \ldots, w$; which suffice to determine the same line yz.

This instance easily leads to the Clebsch transference principle, which is:

If $$I = f\{(ab), \ldots\}$$

is the symbolic form of an invariant of binary ground forms $a_x{}^p$, $b_x{}^q$, . . . , *then the corresponding expression*

$$f\{(ab \mid xy), \ldots$$

where each bracket factor has been replaced by a second compound always containing the same x, y, *has the same geometrical significance for the points common to loci* $a_x{}^p = 0$, $b_x{}^q = 0, \ldots,$ *and an arbitrary line* xy *in* (n — 1)*-fold space, that the invariant* I *has for points on the line illustrating a binary field.*

Proof.—

With the notation of this section, if the geometrical property of the $p + q + \ldots$ points on a line is given by $f\{(ab), \ldots\} = 0$, then that of the corresponding points on the line xy is given by $f\{(AB), \ldots\} = 0$.

But $$(AB) = a_x b_y - a_y b_x = (ab \mid xy). \quad . \quad . \quad . \quad (8)$$

Hence $$f\{(ab \mid xy) \ldots\} = 0,$$

which proves the theorem.

Covariants.—More generally a binary covariant involving a variable x is transferred by this principle simply by altering bracket factors, as in (7), and leaving inner products unchanged. This procedure is implied in conditions (6) above.

For example, the binary covariant equation

$$(ab)a_x b_x = 0$$

gives the point pair at once harmonic to a_x^2 and b_x^2. Hence on a line xy, the corresponding points are given by

$$(AB)A_\xi B_\xi = 0,$$

or
$$(ab \mid xy)a_X b_X = 0,$$

or
$$(abu \ldots t)a_x b_x = 0,$$

replacing X by x, provided x denotes the variable point on the line $u \ldots t$ cut by the loci a_x^2, b_x^2. This concomitant of course also has a significance for points x in space, which are not on the line; but in that case it has no direct binary relation.

Transference in General.

An invariant property of a lower can always be transferred to a higher field by this extension of symbolic bracket factors.

For we merely have to replace a bracket factor $(a_1 a_2 \ldots a_r)$ of a field of order $r (< n)$ by an rth compound

$$(a_1 a_2 \ldots a_r \mid x_1 x_2 \ldots x_r),$$

or its equivalent outer product

$$(a_1 a_2 \ldots a_r u_1 u_2 \ldots u_{n-r}),$$

and interpret the result in a field of order n. The reader will have no difficulty in supplying a formal proof by the methods of §4 (cf. p. 184, (8)), by starting with r distinct points $x_1, \ldots, x_r$, and r parameters $\xi_1, \ldots, \xi_r$, in place of the previous ξ_1, ξ_2.

EXAMPLES

1. The ternary condition $(abu)a_x b_x = 0$ gives the points x on a line u which are conjugates with regard to two conics a_x^2, b_x^2.

2. If $(ab \mid xy)a_x b_x \equiv (abuv)a_x b_x \equiv (abp)a_x b_x$ vanishes, the points x on the line p are conjugates for each of two quadric surfaces a_x^2, b_x^2.

3. If $(bcu)(cau)(abu) = 0$, a certain trio of conics is cut in involution by the line u. (§4, p. 218.)

4. Interpret $(bcp)(cap)(abp) = 0$ for a line p in threefold space.

5. The envelope of a line which cuts two conics harmonically is a conic.
[For $(abu)^2 = 0$ is quadratic in the line co-ordinates u_1, u_2, u_3.

6. *The totality of lines which cut two quadric surfaces harmonically is the Battaglini quadratic complex* $(\mathrm{abp})^2 = 0$.

We write $(abp)^2 \equiv (abuv)^2 \equiv (ab \mid xy)^2$ where x, y are two points on the line p, and u, v two planes through the line. If y is fixed, x describes a quadric surface. Dually, if v is fixed, the plane u envelops a quadric surface.

6. Projective Properties.

The Clebsch principle gives an instantaneous proof that the vanishing of an invariant of ground form is indeed the algebraic interpretation of a *projective* property (§2, p. 271). For simplicity let us consider projection, in three dimensions, of a figure in a plane u to a plane v from a vertex θ.

Then we call the point x of plane u the projection of x' in the plane v, if points x, x', θ are in line. It is essential that u and v should be distinct planes, neither passing through the point θ.

Now consider a quaternary invariant of any number of points $\theta, x, y, z, t \ldots.$ Let it be

$$f\{(xyzt),\ (xyz\theta), \ldots\}.$$

If all the points except θ are in the plane u, then $(xyzt)$ vanishes identically, because every four coplanar points are linearly related. Hence the function is entirely composed of factors, each including θ, which we now write $f\{(xyz\theta), \ldots\}$.

But for any other point x' in the line θx we can take

$$x' = x + \lambda\theta$$

(λ scalar). Hence, if also $y' = y + \mu\theta$, $z' = z + \nu\theta$, then

$$(x'y'z'\theta) = (x + \lambda\theta \,.\, y + \mu\theta \,.\, z + \nu\theta \,.\, \theta) = (xyz\theta),$$

since all other terms involve two or more θ's in the expansion of the bracket factor, and consequently vanish. Thus

$$f\{(xyz\theta), \ldots\} = f\{(x'y'z'\theta), \ldots\}$$

showing that actual projection maintains the invariant property.

Further, if we suppress θ in each factor, the invariant $f\{(xyz), \ldots\}$ now exhibits a *ternary* property of points in the plane u. It follows that every ternary invariant equated to zero specifies a projective property. Such an argument is general, true for all dimensions, for any number of successive projections; and indeed can be extended to include symbols as well as points among the bracket factors.

7. First Geometrical Interpretation of Linear Transformation. Collineation.

Let M be the non-singular square matrix of n rows, and x a single-column matrix representing a point P in $(n-1)$-fold space. Then the product Mx is also a single column, which denotes a point Q. We write

$$M=[e_{ij}], \qquad |M|=|e_{ij}|\neq 0, \qquad x=\{x_i\}, \qquad \xi=\{\xi_i\},$$
$$Mx=\xi, \qquad M^{-1}\xi=x.$$

We use homogeneous co-ordinates, so that, if $\rho\neq 0$, $\{\rho x_1, \ldots, \rho x_n\}$ represents the same point P. If x is given, a unique set ξ is found, and if ξ is given, x also is unique. Also since

$$\rho Mx=M\rho x$$

when ρ is scalar, the point P given by the set $\{\rho x_1, \ldots, \rho x_n\}$ in this way is connected with the point Q, $\{\rho\xi_1, \ldots, \rho\xi_n\}$ by what is called a *one-one correspondence.* Given either point P, Q the other is completely determined.

Again if the matrix M is replaced by any other, except a mere multiple ρM of itself, a new point Q is derived from the same point P. So the correspondence between P and Q is specified by the matrix.

Geometrically, when a one-one correspondence connects points P of a given field with points Q of a second given field (which may coincide with the field of P, as in the present case), the correspondence is called a *collineation.* Thus:

For a given frame of reference a non-singular matrix M *of order* n *determines a collineation for points of the field.*

Or again,

A linear transformation x → ξ *determines a collineation between points* P(x) *and points* Q(ξ).

EXAMPLES

1. The equations

$$\begin{aligned}\xi_1&=l_1x_1+m_1x_2+n_1x_3,\\ \xi_2&=l_2x_1+m_2x_2+n_2x_3, \qquad |l_1m_2n_3|\neq 0,\\ \xi_3&=l_3x_1+m_3x_2+n_3x_3,\end{aligned}$$

determine a collineation between the point $P(x_1, x_2, x_3)$ and the point $Q(\xi_1, \xi_2, \xi_3)$ of a plane.

2. If P lies on a given line, Q also lies on a given line.

3. If P describes a curve of order n, so does Q.

4. In ordinary space, if P lies on a given line, so does Q; if P lies on a given plane, so does Q.

5. Generalize this set of results.

6. If $P_1P_2P_3P_4$ are four points in line, the cross ratio $\{P_1P_2P_3P_4\}$ is the same as that of the corresponding points $\{Q_1Q_2Q_3Q_4\}$.

7. An involution on a straight line (§4, p. 218) is a *symmetrical* collineation. Thus if, on a line Λ, P corresponds to Q then when P is situated at Q, Q becomes P.

8. The general definition of involution in space is *symmetry* of collineation. Prove that the necessary and sufficient condition for an involution is $M^2 = \rho I$, where ρ is any non-zero scalar.

[$M^{-1} = \rho M$.

8. Latent Points of a Transformation.

For certain positions of $P\{x_1, x_2, \ldots, x_n\}$, the corresponding points P and Q will coincide. These are called latent points.

In general there will be n distinct latent points; for if P and Q coincide, we have for some value λ,

$$e_{r1}x_1 + e_{r2}x_2 + \ldots + e_{rn}x_n = \lambda x_r, \qquad r = 1, 2, \ldots, n. \quad (9)$$

Written in full this gives n linear equations from which, by eliminating $x_1, x_2, \ldots, x_n$, we have the so-called characteristic equation (p. 98) of the matrix:

$$f(\lambda) \equiv \begin{matrix} e_{11}-\lambda & e_{12} & & e_{1n} \\ e_{21} & e_{22}-\lambda & & e_{2n} \\ & & & \\ e_{n1} & e_{n2} & & e_{nn}-\lambda \end{matrix} = 0. \quad (10)$$

In the case when there are n distinct roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ of this equation, which is a binary n-ic in λ, we determine one set of values $x^{(i)}$ by solving $(n-1)$ of the equations (9) for each value λ_i of λ. Thus we should have $x^{(i)}$ given by a row of first minors of the determinant $f(\lambda_i)$, and the n sets $x^{(i)}$ so found will be distinct. For if not, let $\rho x^{(i)} = \sigma x^{(j)}$. Then substituting $x^{(i)}$ and $x^{(j)}$ in turn in (9) and subtracting, $\rho\lambda_i x^{(i)}{}_r = \sigma\lambda_j x^{(j)}{}_r$, i.e. $\lambda_i = \lambda_j$, which is contrary to the hypothesis. These n points are also linearly independent and form a simplex (p. 296).

The particular simplex (p. 86) whose n primes are given by the n equations $x_i = 0$ is called the *frame of reference.* If $n = 3$, it is the familiar triangle of reference; if $n = 4$, the tetrahedron of reference; and so on.

Example.—

A collineation referred to its latent points as frame of reference takes the form $x_1 = \lambda_1 \xi_1$, $x_2 = \lambda_2 \xi_2$, . . . , $x_n = \lambda_n \xi_n$.

9. Second Geometrical Interpretation of Linear Transformation. Change of Frame of Reference.

Instead of maintaining a fixed frame of reference and interpreting a linear transformation as a collineation, we may consider that the geometrical figure is fixed, but a change is made in the frame of reference, exactly as was done in Ex. **1**, p. 151. Again it will be found that this illustrates the same algebra.

The three homogeneous point co-ordinates x_i of that example are replaced by n such co-ordinates; and the three line co-ordinates u_i, by n prime co-ordinates u_i. Then $u_x = 0$ is the point equation of the prime u (or dually is the prime, or tangential, equation of the point x). Accordingly, we interpret contragredient linear transformations $x \rightarrow x'$, $u \rightarrow u'$, for which $u_x = u'_{x'}$, as giving the same point x and the same prime u referred to a new frame.

Let the point P referred to one simplex have co-ordinates $x = \{ x_1, x_2, \ldots, x_n \}$, and to another have co-ordinates $x' = \{ x_1', x_2', \ldots, x_n' \}$.

Also let

$$x_r = c_{r1}x_1' + c_{r2}x_2' + \ldots + c_{rn}x_n', \quad r = 1, 2, \ldots, n, \qquad (11)$$

where the matrix $C = [c_{ij}]$ has rank n, so that $| C | \neq 0$. Then the n primes given by $x_r = 0$ have, for equations in terms of $x_1', \ldots, x_n'$,

$$c_{r1}x_1' + \ldots + c_{rn}x_n' = 0.$$

These will be the equations of the primes of the simplex of reference for x in terms of x'. And if we solve equations (11) for x' we likewise obtain the primes of the second simplex referred to the first.

It is well to have these two distinct interpretations of a linear transformation, for both have their value.

10. Reciprocation and Correlation.

It has been remarked (§2, p. 283) that there is a second type of duality besides the reciprocity generated by the texture of space itself. In the second, a certain geometrical locus or manifold or ground form Γ is required which gives rise to the reciprocity. Let us confine our investigation to the case when Γ is given algebraically by a matrix of a bilinear form Φ; and for shortness let it be a ternary form, as typical of the general case.

We consider

$$\Gamma = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = [a_{ij}] \quad . \quad . \quad . \qquad (12)$$

$$\begin{aligned} \Phi = {} & a_{11}x_1y_1 + a_{12}x_1y_2 + a_{13}x_1y_3 \\ & + a_{21}x_2y_1 + a_{22}x_2y_2 + a_{23}x_2y_3 \\ & + a_{31}x_3y_1 + a_{32}x_3y_2 + a_{33}x_3y_3 \\ = {} & \Sigma a_{ij}x_iy_j = a_xb_y \end{aligned} \qquad (13)$$

where $a_{ij} = a_ib_j$, symbolically.

Let x, y denote two points of which y is fixed. Then $\Phi = 0$ gives the equation of a straight line, since it is linear in x. This is called the *polar* line of the point y with regard to Φ, and correlatively y is called the *pole* of this line. Thus if $u_x = 0$ is the polar of y for this bilinear form, then on comparing coefficients we may take

$$\begin{aligned} u_1 &= a_{11}y_1 + a_{12}y_2 + a_{13}y_3 \\ u_2 &= a_{21}y_1 + a_{22}y_2 + a_{23}y_3 \\ u_3 &= a_{31}y_1 + a_{32}y_2 + a_{33}y_3 \end{aligned} \qquad (14)$$

Further, if the determinant $|a_{ij}| \neq 0$, we have by solving these equations,

$$\begin{aligned} y_1 &= a^{11}u_1 + a^{21}u_2 + a^{31}u_3 \\ y_2 &= a^{12}u_1 + a^{22}u_2 + a^{32}u_3 \\ y_3 &= a^{13}u_1 + a^{23}u_2 + a^{33}u_3 \end{aligned} \qquad (15)$$

where a^{ij} is the typical element of the reciprocal $|a^{ij}|$ of the determinant $|a_{ij}|$. Thus

$$\Gamma^{-1} = \begin{array}{ccc} & a^{21} & a^{31} \\ a^{12} & & \\ a^{13} & a^{23} & \end{array} = [a^{ji}] . \qquad (16)$$

The linearity of these conditions (14), (15) shows that a one-to-one correspondence exists between pole and polar for the ground form Γ. Every point of the plane has its polar, and every polar has its pole.

11. Correlation.

An interesting algebraic feature has presented itself in the system of equations (14), (15), connecting *contragredient* variables y and u, for hitherto we have only dealt with cogredient transformations. So we make the following definitions.

Definition of Correlation and Collineation.—*A linear transformation connecting contragredient variables is a correlation, one connecting cogredient variables is a collineation.*

Transformations $x \longrightarrow u$, $u \longrightarrow x$ are correlations, while $x \longrightarrow x'$, $u \longrightarrow u'$ are collineations. We classify correlations as symmetrical, skew symmetrical, and general.

EXAMPLES

1. If Γ is a symmetrical correlation, then its matrix furnishes the coefficients of a quadric, $\Sigma a_{ij} x_i x_j, a_{ij} = a_{ji}$. The correlation $x \longrightarrow u$ is now one aspect of *polar reciprocation* with regard to the quadric: it replaces a point x by its polar.

The inverse correlation $u \longrightarrow x$ replaces a polar by its pole.

2. If Γ is skew symmetrical, then $a_{ij} = -a_{ji}$. Symbolically $a_x b_y = -a_y b_x = \frac{1}{2}(ab \mid xy)$.

Prove that every point lies on its polar. What is the inverse property?

When pole and polar are so *incident* the correlation is called a *Null System.*

3. A quaternary linear complex $(ab\pi_2) \equiv \Sigma a_{ij} p_{ij}$ generates a Null System. For if the line $p = xy$ belongs to the complex then $\Sigma a_{ij} p_{ij} = 0$. If y is fixed, x describes a plane whose equation is $(ab \mid xy) = 0$. This is called the polar plane, and clearly y lies on it.

4. What is the dual of Ex. 3?

5. Can this be generalized?

6. Show that a Null System breaks down in space of even dimensions.

[If n is odd, the skew symmetric matrix Γ is singular.

12. Canonical Form of a Matrix.

A very interesting application of the twofold geometrical interpretation of a linear transformation is provided by the following theorem.

A matrix M *all of whose latent roots* $\lambda_1, \lambda_2, \ldots, \lambda_n$ *are distinct from each other can be expressed as a product* ALA^{-1}, *where* A *is non-singular and* L *is a diagonal matrix consisting of* $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Proof.—

For let M denote the collineation changing a point x to ξ, so that $x = M\xi$. Since M has n distinct latent roots it has a simplex of n latent points. For, if there were a relation

$$\rho_1 x^{(1)} + \rho_2 x^{(2)} + \ldots + \rho_k x^{(k)} = 0, \ldots \qquad (17)$$

with *none* of the ρ's zero, for k of the latent points $x^{(1)}, \ldots, x^{(k)}$ $(k \leqslant n)$, then, as in the special case of §8, we should have $\Sigma \rho_i \lambda_i x^{(i)} = 0$. From this and (17) one $x^{(i)}$ could be eliminated, giving a relation for $k - 1$ points, then similarly for $k - 2$ points, and finally for one point, which is absurd. The k points are therefore linearly independent.

Let the change of frame from the original to this new simplex be given by $x = Ay$ and $\xi = A\eta$, so that the cogredient sets x, ξ are now replaced by y, η.

Hence $y \longrightarrow \eta$ is the collineation referred to its latent points as frame. This changes y_1 to $\lambda_1\eta_1$, y_2 to $\lambda_2\eta_2, \ldots, y_n$ to $\lambda_n\eta_n$ (Example of §8, p. 293), so that $y = L\eta$ where L is the diagonal matrix of $\lambda_1, \lambda_2, \ldots, \lambda_n$.

But by elimination of η, y we have

$$M\xi = x = Ay = AL\eta = ALA^{-1}\xi$$

true for every point ξ. Hence $M = ALA^{-1}$. This proof is not invalidated if one of the latent roots of M is zero.

The corresponding theorem[1] when latent roots are coincident is true, provided L is suitably modified.

EXAMPLES

1. By expanding the equation $MA = AL$ in the above, for the case of three-rowed square matrices, verify that $\lambda_1, \lambda_2, \lambda_3$ are latent roots of the matrix M, and that there are nine linear equations, from which the elements of M can actually be determined.

2. Prove the Cayley Hamilton theorem by using the canonical form of a matrix.

---

[1] Cullis, *Matrices and Determinoids* (Cambridge, 1926), III, p. 342. Dickson, *Modern Algebraic Theories* (Chicago, 1926), Chap. V. Bôcher, *Higher Algebra* (New York, 1919), Chap. XX. Turnbull and Aitken, *Canonical Matrices* (Glasgow, 1932).

CHAPTER XX

THE GENERAL QUADRIC

1. Complete System of the General Quadric.

Let $$f = a_x^2 = b_x^2 = c_x^2 = \dots = m_x^2$$

be the symbolic form of the quadric $F = \Sigma a_{ij} x_i x_j$ homogeneous in n variables $x_1, x_2, \dots, x_n$, $(n > 2)$, where $a, b, \dots, m$ are n equivalent symbols, defined by the identity

$$a_{ij} = a_{ji} = a_i a_j = b_i b_j = c_i c_j = \dots = m_i m_j,$$

true for all values of i and j from 1 to n inclusive.

We can prove that *the symbolic expressions*

$$u_x,\ f = a_x^2,\ \Delta_2 = (ab \mid xy)^2,\ \Delta_3 = (abc \mid xyz)^2, \dots$$
$$\Delta_{n-1} = (ab \dots l \mid xy \dots)^2,\ \Delta = (ab \dots lm)^2,$$

all of which are concomitants of the quadric, form a complete irreducible system for a single quadric $\mathrm{f} = \Delta_1 = \mathrm{a}_\mathrm{x}^2$ *and any number of linear ground forms.* In other words, every polynomial concomitant of one quadric and any number of variables $x, y, z, \dots$, $u, v, w, \dots$ is expressible as a polynomial in $u_x, \Delta_1, \Delta_2, \dots, \Delta_n$ and of polars of these forms.

Proof.—

Let each variable $x, y, \dots$ be resolved into $(n - 1)$ cogredient variables of type u, so that any concomitant is now symbolized by a polynomial in outer products only, such as

$$(A_r U_{n-r}) = (abc \dots uvw \dots).$$

Here there are r equivalent quadric symbols convolved, together with $n - r$ variables $u, v, \dots$.

Now let a single-term symbolic product of ν factors, involved in a concomitant, be

$$P = (A_{r_1} U_{n-r_1})(B_{r_2} V_{n-r_2}) \dots (K_{r_\nu} W_{n-r_\nu}),$$

where the suffixes $r_1, r_2, \dots, r_\nu$ are arranged in descending order. Somewhere in the product beyond the first factor will be found the r_1 duplicate symbols of A_{r_1}. Either they are all in the second factor or not. If they are, then $r_2 = r_1$ since r_2 cannot exceed r_1, and the first two factors of P are

$$\Delta_{r_1}' = (A_{r_1} U_{n-r_1})(A_{r_1} V_{n-r_1}),$$

an actual polarized form of Δ_{r_1} or Δ_{r_1} itself. Then P is said to be reducible; we remove this factor and deal with the residual lower degree factor.

But if B_{r_2} does not contain all the r_1 duplicates of A_{r_1}, we transform P by the process of §8, p. 193, and convolve A_{r_1} a second time in this second factor at the expense of other symbols and variables originally within the factor. Thus

$$P = \Sigma \lambda (A_{r_1} U_{n-r_1})(A_{r_1} B_i' V_j') \dots,$$

where λ is numerical, and $B_i' V_j'$ is part of the original contents of this factor. If $i = 0$, again a factor Δ_{r_1}' emerges. If $i > 0$ we place this second factor first, with its increased currency $r_1 + i$ of equivalent symbols, and proceed as before.

Since i cannot exceed $n - r_1$, the process of so raising the currency is finite, and P is thereby expressed as reducible terms containing factors Δ_r, $r = 0, 1, \dots, n$. This proves the theorem.

Corollary I.—*A single quadric has only one invariant—its discriminant.*

Corollary II. — *The complete system for the dual form* $\Sigma = u_\alpha^2 = u_\beta^2 = \dots$, *is*

$$u_x,\ \Sigma,\ (\alpha\beta \mid uv)^2,\ (\alpha\beta\gamma \mid uvw)^2,\ \dots,\ (\alpha\beta\gamma \dots \mu)^2.$$

EXAMPLES

1. The system for a ternary quadratic $a_x^2 = b_x^2 = c_x^2$ is a_x^2, $(abu)^2$, $(abc)^2$. What is their geometrical significance?

2. A covariant conic exists for a conic and a single point.

Ans. $(ab \mid xy)^2$ is the covariant, if y is the given point, and x the variable.

3. A contravariant conic exists for a conic, in tangential co-ordinates, and a single line.

Ans. $(\alpha\beta \mid uc)^2$ where $u_\alpha{}^2 = u_\beta{}^2$ is the quadratic and c_x the linear form.

4. If $l_x = 0$ is the line at infinity what does $(\alpha\beta \mid lu)^2 = 0$ represent?

[The line is parallel to an asymptote of the conic.

5. *The bordered determinants* (pp. 103–105) *give the non-symbolic form of the irreducible concomitants* Δ_i.

6. *The equations of these concomitants* Δ_i *are the respective conditions that a line* xy, *plane* xyz, . . . *should touch a quadric.*

[Use the Clebsch transference principle.

7. *If* $\Delta = 0$ *the quadric is a cone.*

For if $a_x a_\xi = 0$, $b_y b_\xi = 0, \dots, m_t m_\xi = 0$ are n equations of rank $r = n - 1$, where $(xyz \dots t) \neq 0$, they can be solved for ξ (§9, p. 195). Also, by eliminating ξ, they give $\Delta = 0$. *Any* point θ is given linearly by n general points $x, y, \dots, t$: whence $a_\theta a_\xi = 0$, in particular if $\theta = \xi$, $a_\xi{}^2 = 0$ so ξ is on the quadric. Then if θ is also on the quadric, so again is $\xi + \lambda\theta$ (§4, p. 286), and therefore a line $\xi\theta$ is on the quadric. This identifies ξ as *vertex* and $\xi\theta$ as *generator* of a cone.

If the rank is $n - 2$, ξ lies on a line vertex: when $n = 4$ the cone is now two planes. If the rank is $n - 3$, ξ lies on a plane. And so on.

8. Taking $\Delta = 0$, $r = n - 1$ and the vertex ξ as $\{1, 0, 0, \dots, 0\}$ in variables $X_1, X_2, \dots, X_n$, prove that the quadric must be a function of $X_2, \dots, X_n$ only.

9. If $r = n - 2$, prove that $\Delta = \Delta_{n-1} = 0$ identically. Taking the line vertex as $X_1 = X_2 = 0$ prove the quadric is a function of $X_3, \dots, X_n$ only.

10. If $\Delta_{n-1} = (abc \dots lu)^2 = 0$ but not identically, then the prime u touches the quadric. This Δ_{n-1} can also be written $u_\alpha{}^2$, where $\alpha = abc \dots l$ in the notation of (35), p. 47.

[Cf. p. 287, §4, Ex. **5, 6, 7.**

2. Self-conjugate Simplex.

A triangle xyz is self-conjugate for a coplanar conic if x is pole of yz, y of zx, and z of xy.

A tetrahedron is self-conjugate for a quadric surface if x is pole of a plane yzt, y is pole of zxt, z of xyt, and t of xyz. This property can be extended to the general case.

Let

$$\left.\begin{array}{c} x, y, z, \dots, t \\ u, v, w, \dots, q \end{array}\right\} \quad . \quad . \quad . \quad . \quad . \quad (1)$$

denote the n points and corresponding primes of a self-conjugate simplex (p. 86), such that pole and polar are in a vertical column. Then we have the relations (cf. Ex. **4, 8**, p. 287)

$$a_x{}^2 \neq 0, \qquad u_\alpha{}^2 \neq 0, \qquad a_x a_y = 0, \qquad u_\alpha v_\alpha = 0 \quad . \quad (2)$$

as typical of any of the points and primes of the table.

3. Canonical Form of the Quadric.

We can derive an important theorem from this set of relations (2), whereby a general quadric is expressed as the sum of at most n squares of linear forms.

For consider the identity

$$(uvw\ldots q)a_\xi = (avw\ldots q)u_\xi + (uaw\ldots q)v_\xi + \ldots + (uvw\ldots a)q_\xi.$$

By utilizing the self-conjugate property of the simplex (1) we write

$$(avw\ldots q) = a_x, \qquad (uaw\ldots q) = a_y, \quad \&c.$$

Thus

$$(uvw\ldots q)a_\xi = a_x u_\xi + a_y v_\xi + \ldots + a_t q_\xi,$$

which is true for all values of a. By squaring this identity we obtain

$$(uvw\ldots q)^2 a_\xi^2 = a_x^2 u_\xi^2 + a_y^2 v_\xi^2 + \ldots + a_t^2 q_\xi^2,$$

a result of fundamental importance. All the product terms on the right have disappeared because of the conjugate properties such as $a_x a_y = 0$.

Now regarding $x, y, \ldots, t, u, \ldots, q$ as constant, and ξ as variable, *we have thus expressed a general quadric*

$$f(\xi) = a_\xi^2 = \Sigma\, a_{ij}\xi_i\xi_j$$

as the sum of n *squares*

$$F = A_1X_1^2 + A_2X_2^2 + \ldots + A_nX_n^2,$$

where

$$\frac{A_1}{f(x)} = \frac{A_2}{f(y)} = \ldots = \frac{A_n}{f(t)} = \frac{1}{(uvw\ldots q)^2},$$

and where

$$X_1 = u_\xi, \quad X_2 = v_\xi, \ldots, X_n = q_\xi$$

are linear forms in the original variables.

This is called a normal or canonical form of the quadric. It has a very simple matrix of coefficients, consisting of diagonal elements $A_1, A_2, \ldots, A_n$ only. In making this reduction we had all $(n-1)$-fold space, except on the quadric itself for the choice of the point x: one less dimension for the choice of y: and so on. Algebraically we sum this up by saying that the reduction to canonical form is possible in ∞^N ways, where

$$N = (n-1) + (n-2) + \ldots + 2 + 1 + 0 = \tfrac{1}{2}n(n-1).$$

EXAMPLES

1. Use the dual identity

$$(xyz \ldots t)^2\theta_\alpha^2 = u_\alpha^2\theta_x^2 + v_\alpha\ \theta_y^2 + \ldots + q_\alpha^2\theta_t^2$$

to reduce the tangential form of the general contravariant quadric also to the sum of n squares, $\Sigma\, B_iU_i^2$.

2. If $U_1, U_2, \ldots, U_n$ are contragredient to $X_1, X_2, \ldots, X_n$ and in fact denote the same self-conjugate simplex, then the dual form of F is

$$\Sigma = \frac{U_1^2}{A_1} + \frac{U_2^2}{A_2} + \ldots + \frac{U_n^2}{A_n}.$$

This follows by direct calculation from the bordered determinant of F.

3. Find canonical forms for all members of the complete system.

Each member Δ_i is a sum of squares of ith compound co-ordinates: while $\Delta = A_1A_2 \ldots A_n$.

4. If the rank of $[a_{ij}]$ is r, then $\Delta_k \equiv 0$, $k > r$ but Δ_r does not vanish. Prove that the canonical quadric is now the sum of r squares.

4. Theory of Two Quadrics.

Let

$$\left.\begin{aligned} f &= a_x^2 = b_x^2 = c_x^2 = \ldots \\ f' &= r_x^2 = s_x^2 = t_x^2 = \ldots \end{aligned}\right\} \qquad (3)$$

be the symbolic forms of two different quadrics

$$F = \Sigma\, a_{ij}\, x_i\, x_j, \qquad F' = \Sigma\, r_{ij}\, x_i\, x_j \ldots \qquad (4)$$

in n variables. From these two we derive a new quadric $\lambda F + \lambda' F'$, said to belong to the pencil of quadrics determined by F and F'. Geometrically, whatever is common to the quadric manifolds F and F' is common to each of the ∞^1 members of the pencil. For example, two conics have four points in common, shared also by the members of their pencil, when $n = 3$; two quadric surfaces have a curve in common, when $n = 4$; and so on.

Now since the typical coefficient of the quadric $\lambda F + \lambda' F'$ is $\lambda a_{ij} + \lambda' r_{ij}$, the discriminant must be $| \lambda a_{ij} + \lambda' r_{ij} |$, which on expansion is a binary n-ic in $\lambda : \lambda'$, say

$$\delta_\lambda^n = \Delta\lambda^n + \Theta_1\lambda^{n-1}\lambda' + \ldots + \Theta_i\lambda^{n-1}\lambda'^i + \ldots + \Delta'\lambda'^n. \quad (5)$$

In ternary forms we generally write this as

$$\begin{array}{lll} \lambda a_{11} + \lambda' r_{11}, & \lambda a_{12} + \lambda' r_{12}, & \lambda a_{13} + \lambda' r_{13} \\ \lambda a_{21} + \lambda' r_{21}, & \lambda a_{22} + \lambda' r_{22}, & \lambda a_{23} + \lambda' r_{23} \\ \lambda a_{31} + \lambda' r_{31}, & \lambda a_{32} + \lambda' r_{32}, & \lambda a_{33} + \lambda' r_{33} \end{array}$$

$$= \Delta\lambda^3 + \Theta\lambda^2\lambda' + \Theta'\lambda\lambda'^2 + \Delta'\lambda'^3. \quad . \quad . \quad (6)$$

Manifestly Δ and Δ' are the discriminants of F and F' respectively, while the $n - 1$ intermediate coefficients Θ_i are simultaneous invariants derived in succession by repeated application to Δ of the Aronhold operator

$$\left(r\frac{\partial}{\partial a}\right) = \Sigma\, r_{ij}\frac{\partial}{\partial a_{ij}},$$

summed for the $\frac{1}{2}n(n + 1)$ effectively distinct coefficient suffixes ij. It is also clear for geometrical reasons that the ratios $\Delta : \Theta_1 : \ldots : \Delta'$ are absolute invariants, since the condition that the quadric $\lambda F + \lambda' F'$ should degenerate is independent of particular co-ordinate axes. The n roots of the equation in $\lambda : \lambda'$, obtained by equating (5) to zero, are in fact examples of irrational invariants of F and F'.

Probably the reader is familiar with these four invariants Δ, Θ, Θ', Δ' of two conics[1] as they provide interesting properties of the usual analytical geometry. A relation involving them expresses a geometrical fact about the conics, as, for example, that $\Theta^2 = 4\Delta\Theta'$ if a triangle can be inscribed in the conic F' which shall circumscribe conic F; or that Θ vanishes if a triangle inscribed in F' is self-conjugate to F.

5. Reduction of Two Quadrics to the Form

$$F = X_1{}^2 + X_2{}^2 + \ldots + X_n{}^2, \quad F' = A_1X_1{}^2 + A_2X_2{}^2 + \ldots + A_nX_n{}^2.$$

Let the linear transformation be given by

$$X_i = \xi_i x_1 + \eta_i x_2 + \ldots + \omega_i x_n.$$

Then we take as our ν parameters $l_1, l_2, \ldots, l_\nu$ (§9, p. 268), the following $n^2 + n$ $(= \nu)$ quantities:

$$\xi_1, \eta_1, \ldots, \xi_2, \eta_2, \ldots, \xi_n, \eta_n, \ldots, \omega_n, A_1, A_2, \ldots, A_n,$$

in this order.

The number of parameters in each X_i is n, so that F, F' have

---

[1] First developed by Salmon. Cf. Salmon's *Conic Sections*, Sixth Edition, Chap. XVIII. A good elementary account is to be found also in Sommerville's *Analytical Conics* (G. Bell & Sons, 1924), pp. 265–294. It is strange that recent books still omit to mention the crucial fact that these invariants form a *complete* system. The present writer remembers the uneasy feeling he had as a student when first reading this theory, and wondering why this set of four invariants was tacitly assumed to tell the whole story.

$n(n+1)$ parameters between them, if each A_i is independent. This fits the number of coefficients in two *general* quadrics. Corresponding to the equations (12), p. 268, there will be ν conditions which equate functions f_i, of the parameters, respectively to the ν coefficients

$$a_{11},\ a_{12},\ \ldots,\ a_{1n},\ a_{22},\ a_{23},\ \ldots,\ a_{nn},$$
$$b_{11},\ b_{12},\ \ldots,\ b_{1n},\ b_{22},\ b_{23},\ \ldots,\ b_{nn},$$

in this order, of the quadrics, $\Sigma a_{ij}x_ix_j$, $\Sigma b_{ij}x_ix_j$. Also we can solve the requisite $n(n+1)$ equations for the parameters, provided no functional relation $\psi(f)=0$ exists (§9, p. 268). This in turn is non-existent if a determinant $\dfrac{\partial f_i}{\partial l_j}$ of order $n(n+1)$ does not vanish identically. By using the identical transformation, $\xi_1=\eta_2=\ldots=1$, $\xi_2=0$, &c., this determinant is seen to be a non-zero expression, $\pm\Pi(A_i-A_j)$, $i\neq j$: and this justifies the canonical form.

More specifically, if when $n=2$ the 2×3 parameters are the usual ξ_1, η_1, ξ_2, η_2 of $X\rightarrow x$, together with A_1, A_2, then the determinant $\left|\dfrac{\partial f_i}{\partial l_j}\right|$ becomes

$$\Delta=\begin{vmatrix} \xi_1 & \eta_1 & \cdot & A_1\xi_1 & A_1\eta_1 & \cdot \\ \cdot & \xi_1 & \eta_1 & \cdot & A_1\xi_1 & A_1\eta_1 \\ \xi_2 & \eta_2 & \cdot & A_2\xi_2 & A_2\eta_2 & \cdot \\ \cdot & \xi_2 & \eta_2 & \cdot & A_2\xi_2 & A_2\eta_2 \\ \cdot & \cdot & \cdot & \xi_1^2 & 2\xi_1\eta_1 & \eta_1^2 \\ \cdot & \cdot & \cdot & \xi_2^2 & 2\xi_2\eta_2 & \eta_2^2 \end{vmatrix}$$

If $\xi_1=\eta_2=1$, $\xi_2=\eta_1=0$, Δ has only one non-zero element in each of col_1, col_3, row_5, row_6. After expanding by cols_{13}, rows_{56}, then $\begin{vmatrix} \xi_1 & A_1\xi_1 \\ \eta_2 & A_2\eta_2 \end{vmatrix}$ is left, giving (A_2-A_1) alone. This method is quite general. We delete the n last rows and those n of the last $\binom{n+1}{2}$ columns which intersect the rows at ξ_1^2, η_2^2, $\ldots$ ω_n^2; then n of the first columns and their analogous rows; then subtract row_i from row_j for n pairs of suitable suffixes, getting a single unit matrix in the first $\binom{n+1}{2}$ columns.

For the difficult case of specialized quadrics, when this determinant vanishes and this canonical form is not justified, the reader should consult a work on Invariant Factors.[1]

EXAMPLES

1. Two general quadrics have a common self-conjugate simplex.

2. The canonical coefficients A_i are the roots of the characteristic equation $| \lambda a_{ij} - r_{ij} | = 0$.

For the equation is invariantive; hence in the canonical form it is $| \lambda - A_i | = (\lambda - A_1)(\lambda - A_2) \ldots \quad (\lambda - A_n) = 0$.

3. The symmetric functions $1, \Sigma A_i, \Sigma A_i A_j, \ldots, A_1 A_2 \ldots A_n$ are the $n + 1$ irreducible invariants.

4. Two quadrics have at least n quadric contravariants.

The tangential equation of $\lambda F + F' = 0$ in canonical form is $\frac{U_1^2}{\lambda + A_1} + \ldots + \frac{U_n^2}{\lambda + A_n} = 0$, leading to a binary $(n - 1)$-ic for λ. The n coefficients are contravariants.

5. The Jacobian of these contravariants is a contravariant of order n, which has n linear factors if the n coefficients A_i are distinct.

[Prove it for $n = 2, 3, 4$ and then generalize.

6. This Jacobian denotes the common self-conjugate simplex.

7. Reciprocate results, 4, 5 and 6.

6. **Complete System of $(n + 1)$ Invariants.**

The n + 1 *forms* $\Delta, \Theta_1, \ldots, \Theta_{n-1}, \Delta'$ *are a complete irreducible system.*

Proof.—

In fact, let I be a polynomial invariant of the two quadrics. Then by the fundamental theorem it can be expressed as ΣP where P is a product of w bracket factors of the type

$$(A_i R_{n-i}) \equiv (a_1 a_2 \ldots a_i r_1 r_2 \ldots r_{n-i}), \quad . \quad . \quad (7)$$

$i = 0, 1, 2, \ldots, n$. Here there are i equivalent symbols convolved in a matrix A_i of currency i, referring to the first quadric F, and $n - i$ symbols in the matrix R_{n-i} for the second quadric F'.

Let the factors of P be arranged from left to right as far as possible in descending order of currency i. As in §8, p. 193, we can convolve the duplicate symbols of the first factor in the second factor and then if necessary rearrange factors in descending order. Finally, we express a typical product as

---

[1] Cf. Bromwich, *Quadratic Forms* (Cambridge Tract, 1906); Jessop, *Line Complex* (Cambridge, 1913); Dickson, *Modern Algebraic Theories* (Chicago, 1926), 133; Bôcher, *Higher Algebra* (New York, 1919), Chap. XX; Turnbull and Aitken, *Canonical Matrices* (Glasgow, 1932).

$$P=(A_{q_1}\ldots)(A_{q_1}\ldots)(A_{q_2}\ldots)(A_{q_2}\ldots)\ldots(A_{q_\nu}\ldots)(A_{q_\nu}\ldots)(\ldots)\ldots \quad (8)$$

where $$n \geqslant q_1 \geqslant q_2 \geqslant \ldots \quad \geqslant q_\nu \geqslant 0,$$

and *all* the symbols of the first quadric are accounted for among the A's. This product P is now said to be *prepared* for the first quadric. The symbols not yet expressed refer entirely to the second quadric. If now $q_1 = n$, P contains the invariant $(A_n)^2 = (a_1 a_2 \ldots a_n)^2$ as factor, and is reducible.

Similarly if P contains a factor $(b_1 b_2 \ldots b_n)$ composed entirely of symbols of F', it is reducible by convolving the duplicate symbols in a second factor, so that the discriminant $(b_1 b_2 \ldots b_n)^2$ emerges. This only happens if $w > 2\nu$, or if $q_\nu = 0$.

Accordingly, we suppose $w = 2\nu$, $q_\nu > 0$, so that the final factor must contain symbols of both quadrics. We next consider the symbols of the second quadric. Allowing for duplicates in the two final factors we write P more fully as

$$P = (A_{q_1}\ldots)\ldots(A_{q_\nu}B_{k_\nu}C_{s_\nu})(A_{q_\nu}B_{k_\nu}D_{s_\nu}), \quad q_\nu + k_\nu + s_\nu = n,$$

where B, C, D refer to the second quadric, and all the $2s_\nu$ symbols in C and D entirely differ. If $s_\nu = 0$, P contains the factor Θ_{k_ν} and is reducible. So we take $s_\nu > 0$.

Now let the possible forms P be examined in the following order:

(i) By ascending weight w.

(ii) When the weight is the same, in ascending degree in the coefficients of the first quadric, and therefore in ascending value of $\sum_{i=1}^{\nu} q_i$.

(iii) When $w = w'$, $\Sigma q_i = \Sigma q_i'$ for two forms P and P', we examine P before P' if $q_1 = q_1', \ldots, q_i = q_i', q_{i+1} > q'_{i+1}$. The value of i is taken in ascending order.

Further than this the order is immaterial. The effect of such an order is to render any process a reducing process, which shifts a symbol a towards the left out of its own factor. For the resulting form (or forms) can then be *prepared* as in (8), when it will be among those already examined.

Now since $s_\nu > 0$, we can as before convolve the duplicates of the $k_\nu + s_\nu$ symbols B, D in the last factor but one. Reference to the fundamental identity, §11, p. 48, shows that this process either shifts C_{s_ν} entirely, leaving

$$\Theta_{k_\nu+s_\nu} = (A_{q_\nu} B_{k_\nu} D_{s_\nu})^2$$

in place of the two final factors, or else shifts some symbols a of A_{q_ν} to the left. The latter case can only give rise to forms already examined.

This proves that every product P is reducible, with the possible exception of $\Delta, \Theta_1, \dots, \Theta_{n-1}, \Delta'$. In other words, every polynomial invariant of two quadrics is expressible as a polynomial in these $n+1$ invariants.

Finally these are irreducible, because a relation expressing any one Θ_i, say, in terms of the remainder is structurally impossible, as is at once seen by examining the degree in *both* sets of coefficients on the left and right of an assumed identity

$$\Theta_i = \Sigma\, \lambda \Delta^k \Theta_1^{k_1} \dots \Delta'^{k_{n-1}}.$$

This proves the theorem.

EXAMPLES

1. Prove that $(abrs)(abct)(crst)$ vanishes identically.

2. Prove $(bcr)(cas)(abt)(rst) = \frac{1}{6}(abc)^2(rst)^2$.

7. Complete Systems involving Variables.

The complete system for two quadrics and all possible variables $x, \pi_2, \dots, \pi_{n-2}, \pi_{n-1}\ (= u)$ has not been discovered except when $n = 2$, 3, or 4. But it can be demonstrated that the number of covariants involving x alone is $n+1$.

It can be shown that n of the $n+1$ covariants are the n coefficients of λ_1, λ_2 in the binary n-ic obtained by forming the dual point quadratic from the tangential form $\lambda_1\Sigma + \lambda_2\Sigma'$, where

Σ is the bordered determinant $\begin{vmatrix} a_{ij} & u_i \\ u_j & 0 \end{vmatrix}$ (§3, p. 101). These give the quadratic covariants f, f' with $n-2$ intermediate quadratics, as is well known in the ternary case. The remaining covariant is their Jacobian, which represents the common self-conjugate simplex of two quadratics.

Also[1] if $n > 2$, a complete system involving any number of cogredient variables $x, y, z, \ldots$ has, besides $n+1$ invariants and $n+1$ covariants, the $n-1$ functional determinants

$$(A_iR_{n-i})\,(A_i \mid \pi_i)\,(R_{n-i} \mid \pi_{n-i}) \qquad i = 1, 2, \ldots, n-1,$$

together with certain others analogous to the Q forms given overleaf. A system including any number of both types of variable x and u has also been found.[2]

By making $n = 2$ this system becomes the binary system for two quadratics a_x^2, r_x^2 already discussed. In this case the functional determinant $(ar)a_xr_x$ coincides with the simultaneous covariant, making, in all, six irreducible forms.

As in the binary case a syzygy connects the square of the Jacobian covariant with the remaining $2n+1$ invariants and covariants.[3]

As for the case where only variables $u, v, w \ldots$ occur, manifestly a concomitant is expressible symbolically by outer products of type

$$(A_iR_jU_k) \equiv (a_1a_2\ldots a_ir_1r_2\ldots r_ju_1u_2\ldots u_k),$$

where $i+j+k = n$ and the three matrices of symbols refer to the quadrics F, F' and the variables respectively. By the preceding methods it can be shown that all possible irreducible forms are included among the following and their polars:

---

[1] Turnbull and Williamson, *Proc. Royal Soc. Edinburgh*, **45** (1925), 149–165, completed in **50** (1929), 8–25.

[2] *Transactions Cambridge Phil. Soc.*, **21** (1909), 197–240, where one x and any number of u, v are considered; and *Edinburgh*, *loc. cit.* (1929), for any number of each type x or u.

For $n = 3$, the ternary case, the system of two quadratics consists of 20 forms: Gordan, *Math. Annalen*, **19** (1882), 529. See also Grace and Young, *Algebra of Invariants* (1904), pp. 280–287. This has been proved to be strictly irreducible: Van der Waerden, *Amsterdam Ak. Versl.*, **32** (1923), 138–147. For three ternary quadratics Ciamberlini found a system of 128 forms, *Giorn. di mat.* (*Battaglini*), **24** (1886), 141. Of these six are reducible. For four or more ternary quadratics see *Proc. London Math. Soc.* (2), **9** (1910), 81–121.

For $n = 4$, the quaternary case of two quadrics, cf. Gordan, *Math. Annalen*, **56** (1903), 1–48; Turnbull, *Proc. London Math. Soc.* (2), **18** (1919), 69–94. This system has 122 concomitants; Williamson, *Journ. London Math. Soc.*, **4** (1929), 182; proved irreducible, Brahmachari, *Proc. Edinburgh Math. Soc.*, **5** (1937), 41–42, but a further check is needed for possible linear syzygies.

[3] Gilham, *Proc. London Math. Soc.* (2), **20** (1921), p. 326.

$$P_{ij} = (A_i R_j U_k)^2, \quad i, j, k = 0, 1, \ldots, \quad i + j + k = n.$$

$$Q = (A_{i_1} U)(A_{i_1} R_{j_1} U')(R_{j_1} A_{i_2} U'') \ldots (A_{i_\nu} R_{j_\nu} U^{(2\nu-1)})(R_{j_\nu} U^{(2\nu)})$$

$$n > i_1 > i_2 > \ldots > i_\nu, \quad j_1 < j_2 < \ldots < j_\nu < n.$$

Example.—

For the ternary case such a completely irreducible system is

$$F = a_x^2,\ F' = r_x^2,\ (\alpha\rho x)^2,\ a_\alpha^2,\ r_\alpha^2,\ a_\rho^2,\ r_\rho^2,\ a_\rho a_x(\rho x y),\ r_\alpha r_x(\alpha x y),$$

where α, ρ are each of currency two in symbols of their respective quadrics.

Other references to the literature will be found in the *Encyklopädie der Mathematischen Wissenschaften*, III, **3**, **6** (1922), and the earlier *Berichte* by W. F. Meyer. More recently with reference to the cases $n = 4$, $n = 6$, cf. *Proc. London Math. Soc.*, **2**, **25** (1926), 303–327, and *Proc. Roy. Soc. Edinburgh*, **46** (1926), 210–222, **48** (1928), 70–91, **56** (1936), 38–49, 155–162. See also *Proc. Kon. Ned. Akad. Wet. Amsterdam* for many works by Weitzenböck, particularly the long series upon projective differential geometry of ruled surfaces in [4], starting in **43** (1940). W. Saddler: "Polar Properties and Canonical Forms", *Jour. London Math. Soc.*, **16** (1941), 167–172. J. H. Grace: *Jour. London Math. Soc.*, **11** (1936), 20–21, proves the strict irreducibility of certain binary and ternary systems. C. W. Gilham: "On Multiple Binary Forms", *ibid.*, **4** (1929), 166–171, **6** (1931), 203–209.

CHAPTER XXI

MISCELLANEOUS RECENT DEVELOPMENTS

1. Restricted Transformations.

Hitherto we have dealt with the projective invariant theory. It is possible to extend the same methods, recently developed by Weitzenböck[1], to special cases in which the transformations are restricted within a subgroup of the general group (§7, p. 161). An invariant of a subgroup is a function which satisfies the invariant definition for all transformations *within the subgroup*: and the more restricted the group the greater will be the number of possible invariants, because they are required to satisfy fewer conditions.

Consider the non-singular coefficient matrices, where $m = n - 1$,

$$M = \begin{bmatrix} e_{11} & & e_{1n} \\ & \cdot & \cdot \\ e_{n1} & & e_{nn} \end{bmatrix} \qquad M_1 = \begin{bmatrix} e_{11} & \dots & & e_{1n} \\ & & & \\ e_{m1} & \dots & & e_{mn} \\ 0 & \dots & 0 & e_{nn} \end{bmatrix}$$

$$M_0 = \begin{bmatrix} e_{11} & & e_{1m} & 0 \\ & & & \\ e_{m1} & \dots & e_{mm} & 0 \\ 0 & \dots & 0 & e \end{bmatrix} \qquad D = \begin{bmatrix} e_{11} & 0 & & \dots & 0 \\ 0 & e_{22} & 0 & \dots & 0 \\ & & & & \\ 0 & 0 & & \cdot & e_{nn} \end{bmatrix} \qquad (1)$$

$$\rho I = [\rho\delta_{ij}], \quad I = [\delta_{ij}].$$

These are in order, as regards degree of restriction, and each generates a group. For the identical transformation $x = Ix'$, *every* function is an invariant: for the scalar transformation $x = \rho I x'$, every *homogeneous* function is an invariant: for the diagonal transformation $x = Dx'$, every *gradient* (isobaric) function is an invariant: for the general case $x = Mx'$ we have the preceding invariant theory. What can be said of the intermediate orthogonal and affine cases?

---

[1] Cf. *Invariantentheorie*, Chap. IX–XII.

The affine subgroup is given by a matrix M_1. Here we deduce

$$x_n = e_{nn} x_n', \tag{2}$$

while the other variables x_i have general linear transformations. Since $| M_1 | = | e_{11} \ldots e_{mm} | \, e_{nn} \neq 0$, we can suppose x_n, x_n' to be constants and $x_1, \ldots, x_m$, Cartesian co-ordinates in m-fold space. If $m = 2$ this collineation $x \rightarrow x'$ is seen to leave the line at infinity latent; if $m = 3$, the plane at infinity latent. Parallel lines remain parallel after transformation; and these facts are true for all values of $m (= n - 1)$.

The Fundamental Theorem of symbolic methods for affine transformations now runs as follows:

Every polynomial invariant K *of affine transformations for* p *given ground forms* $\mathrm{f}_1, \mathrm{f}_2, \ldots, \mathrm{f}_\mathrm{p}$ *is identical with a projective invariant of the same ground forms together with a certain linear form* $\mathrm{L} = \mathrm{l}_\mathrm{x}$ *latent in the transformations.*

For example, if y, z, t are three coplanar points, then (yzt), l_y, l_z, l_t are projective invariants of a line l and the points. The invariant

$$I = (yzt) / l_y l_z l_t$$

is absolute for affine transformations. Also, if we take $l_x = x_n = x_3$ as in the preceding work, and we write $x_3 = y_3 = z_3 = 1$ with (y_1, y_2), (z_1, z_2), (t_1, t_2) as Cartesian co-ordinates, then $\frac{1}{2} I$ becomes the area of the triangle yzt.

Again the affine theory of a conic rests on the invariants of a quadratic $a_x{}^2 = b_x{}^2$ and a linear form l_x.

Thus $C = (abl)^2$ vanishes if for these Cartesian co-ordinates the conic is a parabola. Non-symbolically

$$C = \begin{vmatrix} a_1 & b_1 & 0 \\ a_2 & b_2 & 0 \\ a_3 & b_3 & 1 \end{vmatrix} = (a_1 b_2 - a_2 b_1)^2 = 2(a_{11} a_{22} - a_{12}{}^2).$$

2. Preparatory Reductions leading to the Proof of the Fundamental Theorem.

First we consider the latent linear form x_n, and write it symbolically as

$$L = (lx) = l_x = l_1 x_1 + l_2 x_2 + \ldots + l_n x_n, \quad . \tag{3}$$

where

$$l_1 = l_2 = \ldots = l_{n-1} = 0, \quad l_n = 1. \quad . \quad . \quad (4)$$

This artifice brings the particular form L into line with the general type of linear form. It follows that

$$(ly) = l_y = y_n, \quad l_z = z_n, \ldots, \quad (l\alpha) = \alpha_n, \ldots,$$

$$(ab \ldots dl) = \begin{vmatrix} a_1 & b_1 & \ldots & d_1 & 0 \\ a_2 & b_2 & \ldots & d_2 & 0 \\ & & & & 1 \end{vmatrix} = (ab \ldots d)_n \quad . \quad (5)$$

where this last suffix n denotes that the determinant has $n - 1$ rows numbered 1, 2, . . . , $n - 1$.

Next, we state the enunciation of the theorem in terms of possible symbolic types: namely, every polynomial invariant K of the affine group M_1 can be symbolically expressed by the factors

$$\begin{matrix}(ab \ldots m), \\ (\alpha\beta \ldots \mu),\end{matrix} \quad (a\alpha) = a_\alpha, \quad (ab \ldots dl) = (ab \ldots d)_n, \quad (l\alpha) = \alpha_n. \quad (6)$$

Here $a, b, \ldots, \alpha, \beta, \ldots$ denote variables or symbols of the ground forms, while l denotes the set (0, 0, . . . , 0, 1).

Thirdly, if the proof holds for linear forms $a_x, b_x, \ldots, u_\alpha, u_\beta, \ldots$ it will hold as before for the general ground form. In fact the symbolic methods hitherto used, together with polarization and the Aronhold process, still continue to be valid in rendering all ground forms multilinear, as these processes have nothing to do with the coefficients e_{ij} of the matrix M_1, which alone has been modified by the affine conditions.

Fourthly, by expressing each linear form u_α as an $(n - 1)$th compound $(ab \ldots d \mid xy \ldots z) = \Sigma \pm a_x b_y \ldots d_z$, we reduce the problem to that of ground forms

$$a_x, \ b_x, \ \ldots \qquad (7)$$

all of one type, whose symbolic invariant types will now be $(ab \ldots m)$, $(ab \ldots dl)$ only, in place of (6). Then invariants may contain groups of $(n - 1)$ symbols owing to implicit convolution of each symbol α. If these symbols α are finally restored they will merely add the other types $(\alpha\beta \ldots \mu)$, a_α, l_α to the list (cf. §9, p. 207).

3. Characteristic Invariant Property.

Since $a_x = a'_{x'}$ as before, the transformation $a' \rightarrow a$ is given by $a' = M_1{}'a$, involving the *transposed* matrix $M_1{}'$. Hence

$$\left.\begin{array}{ll} a_i{}' = e_{1i}a_1 + \ldots + e_{mi}a_m, & i = 1, 2, \ldots, n-1 \\ a_n{}' = e_{1n}a_1 + \ldots + e_{mn}a_m + e_{nn}a_n, & (m = n-1) \end{array}\right\} . \quad (8)$$

Now if $K = K(a, b, \ldots)$ is a polynomial affine invariant of linear forms (7), then the identity

$$K' = K(a', b', \ldots) \equiv \phi(e_{ij})K(a, b, \ldots) \quad (9)$$

holds for all values of $a_i, b_i, \ldots, e_{ij}, \ldots$, when $a', b', \ldots$ are given by (8). The proof of §2, p. 169 will now apply to show that $\phi(e_{ij})$ can only be a polynomial factor of a power of $|M_1|$. But $|M_1|$ has polynomial factors of two kinds only, $\Delta_{n-1} = |e_{11} \ldots e_{mm}|$, and e_{nn}, where Δ_{n-1} is a general determinant in m^2 arguments and therefore irresoluble. We infer

$$\phi(e_{ij}) = \Delta^r_{n-1}\, e_{nn}{}^s, \qquad K' = \Delta^r_{n-1}\,(e_{nn})^s K. \quad . \quad (10)$$

4. Proof of the First Fundamental Theorem.

This last identity can be written

$$K(a', b', \ldots) \equiv \Delta^r_{n-1}\,(e_{nn})^s K(a, b, \ldots). \quad . \quad (11)$$

We have two cases to consider.

Case (1) $K(a, b, \ldots)$ contains no symbol $a_n, b_n, \ldots$ at all with suffix n.

Case (2) $K(a, b, \ldots)$ contains some symbols with suffix n.

Case (1). K is now a *projective* invariant of linear forms such as

$$a_1x_1 + a_2x_2 + \ldots + a_{n-1}x_{n-1}$$

in the field of order $n-1$. For the transformation (8) $a \rightarrow a'$ in this case contains no $a_n{}', b_n{}', \ldots$, so that K' is free from the argument e_{nn}. Hence in (11), $s = 0$ and $K' = \Delta^r_{n-1}K$. Thus K can be symbolized entirely by means of factors of type

$$(ab \ldots d)_n = (ab \ldots dl).$$

Case (2). Here $K(a_1', b_1', \ldots)$ involves $a_n, b_n, \ldots$. By means of the particular affine transformation

$$\begin{aligned} x_i &= x_i' \qquad (i = 1, 2, \ldots, n-1) \\ x_n &= e_{nn} x_n' \qquad (\Delta_{n-1} = 1) \end{aligned} \tag{12}$$

we gather that (11) is satisfied only if K is homogeneous in the quantities $a_n, b_n, \ldots$. Let us therefore write

$$K = K_1 g_1 + K_2 g_2 + \ldots + K_h g_h \qquad (h \geqslant 2), \quad . \tag{13}$$

where each K_i is free from $a_n, b_n, \ldots$; and each g_i is a form of order s in the set $a_n, b_n, \ldots$. Also let the right side of (13) be brought to its lowest terms as a function of $a_k, \ldots$, so that the number h cannot be diminished any further.

Hence, by (11),

$$K_1' g_1' + \ldots + K_h' g_h' \equiv \Delta_{n-1}^r (e_{nn})^s (K_1 g_1 + \ldots + K_h g_h). \tag{14}$$

If we substitute for each $a_n', b_n', \ldots$ in g_i' on the left by means of

$$a_n' = e_{1n} a_1 + e_{2n} a_2 + \ldots + e_{nn} a_n, \text{ \&c.,} \quad . \tag{15}$$

then each g_i' is a polynomial of order s in e_{nn}. Equating the coefficient of $e_{nn}{}^s$ on both sides, we obtain the identity

$$K_1' g_1 + \ldots + K_h' g_h \equiv \Delta_{n-1}^r (K_1 g_1 + \ldots + K_h g_h);$$

whence

$$K_i' = \Delta_{n-1}^r K_i. \quad . \quad . \quad . \quad . \quad . \quad . \tag{16}$$

As in case (1), each K_i is now a projective invariant of the field of order $n - 1$, and thus can be expressed entirely by means of factors $(ab \ldots d)_n$.

Now let ν be the number of symbols $a, b, \ldots$, in K, which is homogeneous in the n elements $a_1, a_2, \ldots, a_n$ of each such set a. Then either $\nu \geqslant n$ or $\nu < n$. If the former, we choose n symbols $a, b, \ldots, m$ and develop K as a Gordan-Capelli series ((22), p. 254)

$$K \doteqdot K_0 + (ab \ldots m) K_1 + \ldots + (ab \ldots m)^\lambda K_\lambda, \quad (\lambda > 0). \tag{17}$$

Again these forms K_i will be affine invariants, satisfying (11), each of which can be dealt with in the same way, if it still contain

n symbols $a, b, \ldots$. Finally we are left to consider the case of at most $n - 1$ symbols in K, so that $\nu < n - 1$ or $\nu = n - 1$.

If $\nu < n - 1$, no factor $(ab \ldots d)_n$ is possible, although by (16) K_i is expressible by such factors. Hence each K_i can only be a constant c_i, and consequently $r = 0$; so that

$$K = c_1 g_1 + c_2 g_2 + \ldots + c_h g_h = g_s, \quad . \quad . \quad (18)$$

where g_s is a form of order s in $a_n, b_n, \ldots$ alone.

Also, if $\nu = n - 1$, then

$$K = (ab \ldots d)_n^r (c_1 g_1 + \ldots + c_h g_h) = (ab \ldots d)_n^r g_s,$$

and we can discard the factor $(ab \ldots d)_n^r$ which is of the desired type, and confine ourselves entirely to g_s.

Finally, it can be proved that g_s vanishes identically. For by (11), we have the identity

$$g_s' = (e_{nn})^s g_s. \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (19)$$

But if we take the particular $n - 1$ sets of values

$$\begin{array}{l} a = 1 \quad 0 \quad 0 \quad \ldots \quad 0 \quad 0 \\ b = 0 \quad 1 \quad 0 \quad \ldots \quad 0 \quad 0 \\ \\ d = 0 \quad 0 \quad 0 \quad \ldots \quad 1 \quad 0, \end{array}$$

then $a_n = b_n = \ldots = 0$, so that $g_s = 0$, while, by (15),

$$g_s' = g_s(e_{1n}, \; e_{2n}, \; \ldots, \; e_{n-1, n}). \quad . \quad . \quad (20)$$

But (19) now shows that g_s' vanishes, so that $g_s(e_{1n}, \ldots, e_{n-1, n})$ vanishes, although its arguments are arbitrary. Hence it vanishes identically. This completes the proof of the First Fundamental Theorem for affine invariants.

5. Consequences of the Theorem.

Since the Fundamental Theorem links affine invariants with projective invariants, by means of the additional linear form $L \equiv l_x$, it follows at once that all the main theorems apply to this restricted case: the Second Fundamental Theorem, and the theorems of Gordan, Hilbert, Clebsch, and Peano. Further, we can imagine, in the preceding proof, that an arbitrary general linear transformation T_0 has been applied to the original variables, expressing them in terms of a set $\xi_1, \xi_2, \ldots, \xi_n$, and in particular that

$$x_n = q_1\xi_1 + q_2\xi_2 + \ldots + q_n\xi_n.$$

Applied to the affine symbolic forms this replaces $(ab\ldots d)_n$ by a type $(ab\ldots dq)$ and a_n by a_q, which are no other than the types $(ab\ldots dl)$ and a_l already utilized. Hence *any* given linear form may be taken as the latent form of the transformation.

Of course, if the x_n co-ordinate is selected as latent, we must note that certain fundamental identities, not of projective type, will arise (cf. Ex. **4**, p. 51), such as

$$(\alpha\beta\gamma)x_3 = (x\beta\gamma)\alpha_3 + (\alpha x\gamma)\beta_3 + (\alpha\beta x)\gamma_3$$

instead of the usual

$$(\alpha\beta\gamma)l_x = (x\beta\gamma)l_\alpha + (\alpha x\gamma)l_\beta + (\alpha\beta x)l_\gamma.$$

Examples.—

1. Examine the restricted transformation $x = M_1'x'$ where M_1' is the transposed matrix of M_1,

It leaves a *point* u_λ latent. By reciprocating the above work, its concomitants are symbolized by types

$$\begin{matrix}(ab\ \ldots\ m),\\ (\alpha\beta\ \ldots\ \mu),\end{matrix}\quad a_\alpha;\quad (\alpha\beta\ldots\delta\lambda) = (\alpha\beta\ldots\delta)_n,\quad a_\lambda.$$

(*Weitzenböck.*)

2. The *affine group with a fixed point* is given by the matrix M_0. Prove that a point u_λ and a prime l_x are both latent; and that the requisite symbolic invariants of this group are

$$\begin{matrix}(ab\ \ldots\ m),\\ (\alpha\beta\ \ldots\ \mu),\end{matrix}\quad a_\alpha,\quad \begin{matrix}(ab\ \ldots\ dl),\\ (\alpha\beta\ \ldots\ \delta\lambda),\end{matrix}\quad \begin{matrix}a_\lambda\\ l_\alpha\end{matrix},$$

together with the absolute invariant l_λ which is purely numerical.

(*Weitzenböck.*)

6. The Orthogonal Group.

A similar theorem holds when a *quadric* is latent. If the quadric is r_x^2, then the projective theory of ν ground forms $f_1, \ldots, f_\nu$ together with r_x^2 is in close touch with the orthogonal invariant theory of the ν ground forms alone.

This theorem lies at the base of an algebraic account of Euclidean, elliptic, or hyperbolic geometry. For instance, in Euclidean geometry, r_x^2 for ternary forms is taken to be $x_1^2 + x_2^2$, but in the other types it is a general ternary quadratic.

The theorem also covers Riemannian geometry where the element of arc is given by $ds^2 = \Sigma g_{ik} dx_i dx_k = g_{dx}{}^2$, as far as metrical properties of small intervals are concerned.

If a, b, α, β are ternary symbols, then invariants of $r_x{}^2$ are composed of types (abc), $(\alpha\beta\gamma)$, a_α, $r_\alpha r_\beta$, &c. For the Euclidean case, if $r_x{}^2 = x_1{}^2 + x_2{}^2$, then $r_\alpha r_\beta = \alpha_1\beta_1 + \alpha_2\beta_2 = (\alpha \mid \beta)$, which accounts for the importance of the inner product of two vectors α and β, when co-ordinate axes are rectangular.

Again if a, b, α, β are quaternary symbols, then invariants of $r_x{}^2$ are composed of types $(abcd)$, $(\alpha\beta\gamma\delta)$, a_α, $r_\alpha r_\beta$, &c., and we shall now have $r_x{}^2 = x_1{}^2 + x_2{}^2 + x_3{}^2$ for the Euclidean case. Once more $r_\alpha r_\beta$ denotes the inner product $(\alpha \mid \beta)$ but it now has three terms $\alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3$. The variables x_1, x_2 (or x_1, x_2, x_3) may be regarded as rectangular Cartesian co-ordinates for plane (or for three dimensional) geometry; and $r_x{}^2$ denotes the circular points at infinity (or the circle at infinity). We shall return to this at p. 324.

This theorem also throws light on elementary analytical solid geometry, where such formulæ appear as $\cos\theta = ll' + mm' + nn'$ for the angle between two straight lines whose direction cosines are given. For orthogonal transformation this is an invariant; in fact it is an inner product of two unit vectors. Likewise the volume of the tetrahedron, three of whose edges are unit vectors, is $\frac{1}{6}(l\,l'l'')$, in terms of an outer product.

It is a commonplace that inner and outer products should so arise, but the invariant theory shows that such products give a *complete* mechanism for dealing with the geometrical entities.

7. Fundamental Theorem of Orthogonal Transformation.

First let us consider this theorem when the latent quadric can be written as

$$(x \mid x) = x_1{}^2 + x_2{}^2 + \ldots + x_n{}^2$$

so that the transformation is orthogonal (§3, p. 152). Further let us confine the discussion to the proper orthogonal case, by which is meant the case when the determinant Δ of the transformation is unity. If $\Delta = -1$ the transformation is called improperly orthogonal. The proof needs two preliminary lemmas.

Lemma I.—*A proper orthogonal transformation exists which transforms a given unit vector* p *into another such vector* q.

Consider the transformation

$$\frac{2(x \mid p+q)}{(p+q \mid p+q)}(p_i+q_i) - x_i = x_i', \qquad (i = 1, 2, \ldots, n),$$

where $(p+q \mid p+q) = \sum_j (p_j+q_j)^2 \neq 0$. Here $x_i'^2$ can easily be calculated in terms of p_i, q_i, x_i, leading to the result

$$(x' \mid x') = (x \mid x).$$

Hence the transformation $x \longrightarrow x'$ is orthogonal. Furthermore if $(p \mid p) = (q \mid q) \neq 0$, we find, when $x = p$, that x' is q. Hence the transformation turns one given unit vector into another: although it may be an improper transformation. If so, we introduce a third vector such that $(p \mid p) = (q \mid q) = (r \mid r)$, $(p+r \mid p+r) \neq 0$, $(q+r \mid q+r) \neq 0$, and apply the corresponding improper transformations $p \longrightarrow r$, $r \longrightarrow q$. Then the product transformation $p \longrightarrow q$ is necessarily proper. This auxiliary r is also needed if $p+q=0$, to cover the case when $(p+q \mid p+q)$ vanishes and the above $x \longrightarrow x'$ does not exist.

Lemma II.—

$$\left(\frac{\partial}{\partial q} \,\middle|\, \frac{\partial}{\partial q}\right)(q \mid q)^\lambda = 2\lambda(n+2\lambda-2)(q \mid q)^{\lambda-1}$$

For
$$\frac{\partial}{\partial q_i}(q \mid q)^\lambda = \lambda(q \mid q)^{\lambda-1} 2q_i,$$

$$\frac{\partial^2}{\partial q_i^2}(q \mid q)^\lambda = \lambda(\lambda-1)(q \mid q)^{\lambda-2} 4q_i^2 + 2\lambda(q \mid q)^{\lambda-1}.$$

Summing for $i = 1, 2, \ldots, n$ the result follows.

Also
$$\left(\frac{\partial}{\partial q} \,\middle|\, \frac{\partial}{\partial q}\right)(q \mid a)(q \mid b) = 2(a \mid b),$$

and
$$\left(\frac{\partial}{\partial q} \,\middle|\, \frac{\partial}{\partial q}\right)(q \mid k)(ab \ldots hq) = 2(ab \ldots hk),$$

and
$$\left(\frac{\partial}{\partial q}\middle|\frac{\partial}{\partial q}\right)(q\,|\,q)\,(q\,|\,a) = (2n+4)\,(q\,|\,a),$$

and
$$\partial q \quad \left.\frac{\partial}{\partial q}\right)(q\,|\,q)\,(ab\,\ldots\,hq) = (2n+4)\,(ab\,\ldots\,hq).$$

8. **First Fundamental Theorem for Proper Orthogonal Invariants.**

Every polynomial invariant of the proper orthogonal group for ground forms $f_1, f_2, \ldots$ *can be symbolized entirely by the use of two kinds of factors,*

$$(ab\,\ldots\,hk), \qquad (a\,|\,b),$$

the outer and inner products respectively.

Proof.—

This follows by induction. For if $n = 1$, the matrix M is the scalar unit, and the proper orthogonal transformation is merely $x = x'$, the identical transformation. Then every vector is its own outer product and the theorem is obvious. So we assume it for m, and set about proving it for $m + 1 = n$.

Consider the transformation coefficient matrices,

$$M_0 = \begin{array}{cccc} e_{11} & \ldots & e_{1m} & 0 \\ & & & \\ & & e_{mm} & \\ & & 0 & \end{array} \qquad M_{00} = \begin{bmatrix} e_{11} & \ldots & e_{1m} \\ & & \\ e_{m1} & & e_{mm} \end{bmatrix} \qquad (21)$$

If M_0 is orthogonal in the field n, so also is M_{00} for the field $n - 1$ $(= m)$, as is apparent by forming inner products of each pair of columns of either. Further if $|\,M_0\,| = 1$ so also is $|\,M_{00}\,|$. Hence they are both properly orthogonal, if either is.

Now M_0 corresponds to the transformation which leaves the vector $q = \{0, 0, \ldots 0, 1\}$ latent. We note that this is a unit vector, since

$$(q\,|\,q) = q_n^2 = 1. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (22)$$

Let γ_0 denote the group of transformations

$$x = M_0 x', \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (23)$$

which transform the first m components $x_1, x_2, \ldots, x_m$ by means of the matrix M_{00}, but leave $x_n = x_n'$ latent. Then any invariant of the group G_0 is an invariant for its subgroup γ_0.

Now if we express any polynomial invariant of the given field. as $\Sigma k_i g_i$ where g_i is a function solely of the components $a_n, b_n \ldots$, and k_i a function of $a_j, b_j \ldots (j \neq n)$, then each k_i will be an orthogonal invariant in the group γ_0. Consequently, by hypothesis, our invariant is a polynomial in two types of factor, which we write

$$(ab \ldots h)_m = |a_1 b_2 \ldots h_m|, \quad (a \mid b)_m = a_1 b_1 + \ldots + a_m b_m, \quad (24)$$

together with the third type, $a_n, b_n \ldots$ of suffix n.

Also by using the unit vector $q = \{0, 0, \ldots 0, 1\}$ we can write these three types as functions of inner and outer products in the higher field of order n; as is at once apparent when each is fully expanded:

$$\left.\begin{aligned} (ab \ldots h)_m &= \frac{(ab \ldots hq)}{\sqrt{q|q}}, \quad (a|b)_m = \frac{(q|q)(a|b) - (q|a)(q|b)}{(q|q)}, \\ \text{and} \qquad {}_{\iota} &= \frac{(q|a)}{\sqrt{q|q}}: \quad b_n = \frac{(q|b)}{\sqrt{q|q}}, \quad \&c. \end{aligned}\right\} \quad (25)$$

Hence every polynomial invariant of the group G_0 is a polynomial function of arguments

$$(q \mid q), \quad (q \mid a), \quad (ab \ldots hq), \quad (a \mid b), \quad . \quad . \quad (26)$$

divided by a positive integral power of $\sqrt{q \mid q}$.

Now $\dfrac{q}{\sqrt{q \mid q}}$ is a unit vector whatever q may be; and, by Lemma I, a proper orthogonal transformation exists which changes any given unit vector into a second. So instead of the special vector $\{0, 0, \ldots 0, 1\}$ we can now take q to be *any arbitrary unit vector*, so long as, instead of the subgroup γ_0, we take the similar subgroup $\gamma_0{}^*$ which leaves the vector q latent.

Also, all these five types, as now written, are unchanged by any proper orthogonal transformation, such as that which replaces the very special unit vector q by any arbitrary unit vector $q/\sqrt{q \mid q}$, none of whose components vanish. Hence a typical invariant is given by

$$I = (G_1 + G_2 \sqrt{q \mid q}) \,/\, (\sqrt{q \mid q})^k, \quad . \quad . \quad (27)$$

where G_1 and G_2 are polynomials of the same type

$$G\{(q \mid q),\ (q \mid a),\ \ldots,\ (ab \ldots hq),\ \ldots,\ (a \mid b), \ldots\},$$

and k is necessarily a positive integer, since one q enters into every type (25).

Now G_1 cannot be zero, else we could cancel out $\sqrt{q \mid q}$ and then treat G_2 as a new G_1. This being so, G_2 must be zero; for otherwise we could always express $\sqrt{q \mid q}$ *rationally* in terms of I, the components q_i, and a, b, . . . , involved in this equation; which is impossible even in the case

$$q = \{0, 0, \ldots 0, 1, 1\}.$$

Hence $G_2 \equiv 0$, so that k must be even, to make the right and left sides agree in rationality. Thus we write

$$(q|q)^{\lambda} I = G\{(q|q),\ (q|a), \ldots,\ (ab \ldots hq),\ \ldots,\ (a|b),\ \ldots\}, \qquad (28)$$

where λ is a positive integer, and each q_i is non-zero.

Operating on both sides of this identity with $\left(\frac{\partial}{\partial q} \middle| \frac{\partial}{\partial q}\right)$, we find, by Lemma II.

$$(q \mid q)^{\lambda-1} I = G' \{(q \mid q),\ \ldots,\ (ab \ldots hk)\}$$

where G' is of the same type as G, but may involve the outer product $(ab \ldots hk)$ which excludes q.

If we proceed λ times, this operation annihilates $(q \mid q)$ on the left, leaving a non-zero multiple of I, and consequently all the q's disappear on the right, leaving only the types $(a \mid b)$, $(ab \ldots hk)$. This proves the theorem.

9. The Hermitian Transformation with an Absolute Quadric.

If we apply a general linear transformation $\xi = Mx$, of matrix $M = [e_{ij}]$, $|M| \neq 0$, to the canonical quadratic

$$(\xi \mid \xi) = \xi_1{}^2 + \xi_2{}^2 + \ldots + \xi_n{}^2, \qquad (29)$$

we obtain a quadratic in $x_1, \ldots, x_n$, which we symbolize by Thus

$$(\xi \mid \xi) = \sum_{i,j,k} e_{ij} e_{ik} x_j x_k = \sum_{j,k} (e_j \mid e_k)\, x_j x_k = r_x{}^2, \qquad (30)$$

where $(e_j | e_k)$ denotes the inner product $e_{1j}e_{1k} + \ldots + e_{nj}e_{nk}$, and all summations run from 1 to n. Hence

$$r_{jk} = r_j r_k = (e_j \mid e_k).$$

Let us find out what becomes of the last theorem when the variables ξ, which enter an orthogonal transformation $\xi \rightarrow \xi'$, now undergo the further transformation $\xi = Mx$. Then, if α, β denote linear sets cogredient with ξ,

$$(\alpha \mid \beta) = \tfrac{1}{2} \left(\alpha \,\middle|\, \frac{\partial}{\partial \beta}\right) (\beta_1{}^2 + \beta_2{}^2 + \ldots + \beta_n{}^2);$$

from which it follows that $(\alpha \mid \beta)$ is a polar form of $(\alpha \mid \alpha)$. Hence

$$(\alpha \mid \alpha) = r_{\alpha'}{}^2, \quad (\alpha \mid \beta) = r_{\alpha'} r_{\beta'}, \quad . \quad . \quad (31)$$

in terms of the corresponding linear sets α', β' after transformation. Now the inner product of two *cogredient* symbols α, β is not a projective invariant, but only arises as an orthogonal invariant. Here, however, we have expressed it as $r_{\alpha'} r_{\beta'}$, a projective invariant of linear symbols together with the symbols r of the quadric. In so doing we link the orthogonal theory with the projective theory. For if all the ground forms of an orthogonal system are symbolized, as may be done, entirely by cogredient symbols $\alpha, \beta, \gamma \ldots$, then their invariants involve two types only,

$$(\alpha\beta \ldots \mu), \quad (\alpha \mid \beta), \qquad (32)$$

of which the former is already a projective invariant, giving

$$(\alpha\beta \ldots \mu) = | M | \, (\alpha'\beta' \ldots \mu'),$$

when $\alpha \rightarrow \alpha'$, $\beta \rightarrow \beta'$, $\ldots$, $\mu \rightarrow \mu'$.

Furthermore, any non-degenerate quadric $r_x{}^2$ may be reduced to the sum of n squares $\Sigma A_i X_i{}^2$ by a suitable linear transformation (§3, p. 300): so that if also $\xi_i = A_i^{\frac{1}{2}} X_i$ $(i = 1, 2, \ldots, n)$, the resultant transformation T: $x \rightarrow \xi$ is still linear. Conversely, by $T^{-1}: \xi \rightarrow x$ we pass from the orthogonal absolute $(\xi \mid \xi)$ to any given non-degenerate quadric $r_x{}^2$, and thereby we solve the Hermitian problem (§5, p. 158) of the restricted transformations x' which leave a given quadric absolutely invariant:

$$r_x{}^2 \equiv r'_{x'}{}^2.$$

For on performing with the transformation T, we reduce it to a proper orthogonal problem. Hence

For ground forms whose symbols α, β, . . . *are cogredient with* x, *every polynomial invariant of the subgroup in which a given non-degenerate quadric* $\mathrm{r}_{\mathrm{x}}^2$ *is an absolute invariant, can be symbolized by two types of factor*

$$(\alpha\beta \dots \mu), \quad r_\alpha r_\beta. \quad . \quad . \quad . \quad . \quad . \quad . \qquad (33)$$

In other words, *proper orthogonal invariants of a system of ground forms* (f) *are projective invariants of the system* (f, $\mathrm{r}_{\mathrm{x}}^2$), *obtained by adjoining the absolute quadric* $\mathrm{r}_{\mathrm{x}}^2$. *Conversely all projective invariants of* (f, $\mathrm{r}_{\mathrm{x}}^2$) *are proper orthogonal invariants of* (f).

Proof of the Converse.—

For starting with the projective system whose symbols are r, α, β, γ . . . , the only types we require by the Fundamental Theorem (§7, p. 203) are

$$(\alpha\beta \dots \mu), \quad r_\alpha, \quad (rs \dots t),$$

where r, s, . . . t are equivalent symbols. But the presence of $(rs \dots t)$ in an invariant implies the discriminant $(rs \dots t)^2$ (§8, p. 194). If this is rejected, any further factors r_α must occur in pairs $r_\alpha r_\beta$. But $(rs \dots t)^2 = n! \mid r_{ij} \mid = n!$, a pure number, in the case when the quadric r_x^2 is $(x \mid x)$. Thus the two types (33) alone are actually necessary, and the theorem is proved.

EXAMPLES

1. The improper orthogonal case is obtained by taking

$$r_x^2 = x_1^2 + \dots + x_{n-1}^2 + x_n^2,$$

and transforming the results of the proper case by the matrix of zeros with a leading diagonal 1, 1, . . . 1, — 1. The required types are $(\alpha\beta \dots \mu)$, $(\alpha \mid \beta)$; but the outer product changes sign after improper orthogonal transformation.

2. The Lorentz transformation leaves

$$r_x^2 = x_1^2 + x_2^2 + x_3^2 - c^2 x_4^2$$

an absolute invariant, where c^2 is a constant.

The invariant types are $(\alpha\beta\gamma\delta)$, together with

$$r_\alpha r_\beta = \alpha_1\beta_1 + \alpha_2\beta_2 + \alpha_3\beta_3 - c^2\alpha_4\beta_4.$$

This becomes an orthogonal group after the transformation

$$x_1 = \xi_1, \quad x_2 = \xi_2, \quad x_3 = \xi_3, \quad x_4 = \frac{1}{ic}\xi_4.$$

3. Linear transformations, $T: x \longrightarrow x'$, satisfying the Lorentz condition can be constructed by the matrix $(I+SQ)/(I-SQ)$, where

$$\begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ & 1 & \cdot & \cdot \\ & & 1 & \end{bmatrix},$$ and S is skew symmetric.

4. If $S = \begin{bmatrix} \cdot & a \\ -a & \cdot \end{bmatrix}$, $Q = \begin{bmatrix} 1 & \cdot \\ \cdot & -c^2 \end{bmatrix}$, and $v = \dfrac{2ac^2}{1+a^2c^2}$ show that the binary matrix $(I+SQ)/(I-SQ)$ transforms variables x, t to x', t' according to the Lorentz formulæ

$$= \frac{x-vt}{\sqrt{1-\frac{v^2}{c^2}}}, \qquad t' = \frac{t-\frac{vx}{c^2}}{\sqrt{1-\frac{v^2}{c^2}}}.$$

[Use (5), p. 69.

5. If a, b, c are contragredient to α, β, γ, the outer product types (abc), $(ab\gamma)$, $(a\beta\gamma)$, $(\alpha\beta\gamma)$ are all *orthogonal* invariants of ternary forms. Expressed as projective invariants of the absolute $r_x{}^2$, the second and third of these must be modified. Thus if $(\alpha \mid \beta) = r_\alpha r_\beta$, then we can prove that

$$(ab\gamma) = (abr)r_\gamma, \quad (a\beta\gamma) = (ars)r_\beta s_\gamma.$$

For if $f(x) = r_x{}^2 = x_1{}^2 + x_2{}^2 + x_3{}^2$, then $\frac{1}{2}(ars)(rs \mid \beta\gamma) = \frac{1}{2}(ars)(r_\beta s_\gamma - r_\gamma s_\beta)$

$$\begin{vmatrix} a_1 & r_1 r_\beta & s_1 s_\gamma \\ a_2 & r_2 r_\beta & s_2 s_\gamma \\ a_3 & r_3 r_\beta & s_3 s_\gamma \end{vmatrix} = \frac{1}{4} \begin{vmatrix} a_1 & \frac{\partial f(\beta)}{\partial \beta_1} & \\ & & \\ & & \end{vmatrix} = (a\beta\gamma).$$

6. If $r_x{}^2 = s_x{}^2 = x_1{}^2 + x_2{}^2 + x_3{}^2 + x_4{}^2$, then $(abrs)(\alpha\beta \mid rs)$ is equivalent to $2(ab\alpha\beta)$.

[Here $(\Sigma(ab)_{ij}(rs)_{kl})\,(\Sigma(\alpha\beta)_{ij}(rs)_{ij}) = \Sigma(ab)_{ij}(\alpha\beta)_{kl}(rs)_{kl}(rs)_{kl}$, since $r_i r_k = 0$, $(i \neq k)$.

7. Prove, by a Laplace development, that when

$$r_x{}^2 = x_1{}^2 + x_2{}^2 + \ldots + x_n{}^2,$$

then the concomitant

$$(ab \ldots h\, r_1 r_2 \ldots r_{n-p})(\alpha\beta \ldots \delta \mid r_1 r_2 \ldots r_{n-p})$$

involving p linear forms $a_x, b_x, \ldots, h_x$; $n-p$ linear forms $u_\alpha, u_\beta, \ldots, u_\delta$, and $n-p$ equivalent symbols $r_1, r_2, \ldots, r_{n-p}$, may be replaced by the outer product

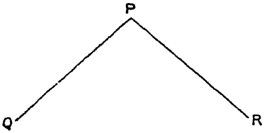
$$(ab \ldots h\, \alpha\beta \ldots \delta)$$

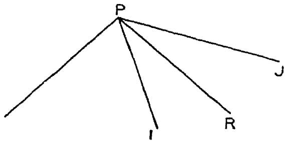
to a numerical factor.

10. Geometrical Significance of the Adjunction Theorem.

The preceding theorems obviously have something in common: they link affine, orthogonal and Hermitian invariants with *projective* invariants by adjoining to given ground forms f, one or more forms ϕ which are latent for the linear transformations of their respective groups. For this reason they are called *Adjunction Theorems.*

If we turn to the geometrical aspect of the theorems we find matter of high interest in metrical geometry. It is well known that properties of distance and size of angle, holding for plane Euclidean geometry, may be interpreted *projectively* by stating them as cross ratio properties of a figure to which the circular points at infinity are adjoined.





Thus, for example, if I, J are the circular points, and P is a point not on IJ, then the pencil $P\{QR, IJ\}$ of four lines through P is harmonic whenever PQ, PR are at right angles.

Analytically, in rectangular Cartesian co-ordinates, the matter is clear if $ax^2 + 2hxy + by^2 = 0$, $x^2 + y^2 = 0$ respectively denote the pairs of lines PQ, PR; PI, PJ. For $a + b = 0$ provided that RPQ is a right angle, or, equally well, if $P\{QR, IJ\} = -1$.

By taking a general conic or quadric, $r_x{}^2$, as latent, we interpret non-Euclidean elliptic or hyperbolic geometry. If $r_x{}^2 = x^2 + y^2$, which is a degenerate conic, we can interpret Euclidean metrical plane geometry, by combining the theorems of §4 and §8. Thus if $m = 2$, $n = 3$, and the matrix M_0 of §1 (1) is orthogonal, then the required result is secured.

EXAMPLE

Thus in ternary symbols, let $l_x = 0$ denote the equation of the line at infinity, and $u_\omega{}^2 = 0$ that of the circular points, so that $u_\omega{}^2$ can be factorized, say

$$u_\omega{}^2 = u_\lambda u_\mu.$$

Then $\{\lambda_1, \lambda_2, \lambda_3\}$ and $\{\mu_1, \mu_2, \mu_3\}$ are the homogeneous co-ordinates of these points I and J.

Furthermore, if $f = a_x{}^2$ denotes a conic whose tangential equation is $u_a{}^2 = 0$, then a covariant conic exists for the quadratics $u_\alpha{}^2$, $u_\omega{}^2$, namely

$$\varphi \equiv (\alpha\omega x)^2 = 0.$$

By the Clebsch principle this gives the locus of a point x whose tangents to these conics form a harmonic pencil. Hence in Euclidean geometry it is the locus of a point whose tangents to a conic are at right angles. In other words the conic φ is the director circle of the conic f.

If this $(\alpha\omega x)^2$ is written down non-symbolically, whether in Cartesian or homogeneous co-ordinates, the ordinary results will be obtained.

Such examples could easily be multiplied, and indeed they form a very attractive analytical projective geometry which has received comparatively little attention.

11. Remarks on the Adjunction Theorem.

It is very tempting to try to discover a general Adjunction Theorem to cover the case when any one or more given ground forms $\phi_1, \phi_2, \ldots, \phi_r$ are latent for a linear transformation T. For if (f) means the complete projective system of concomitants of a set of ground forms f, and (f, ϕ) means that of the whole set of forms f and ϕ, then any member of (f, ϕ) is certainly an invariant of any transformation which leaves each ϕ_i latent. But except for a few cases, detailed above, when ϕ_i is linear or quadratic, *the converse is not true.* Nor has any general law been found to determine restricted transformations for a given set of latent forms ϕ. How this converse applies is still an unsolved problem of the theory.[1]

It is possible to extend these methods of Study and Weitzenböck to the case when a bilinear form is latent, and to the theory of double binary and other multiple fields (p. 240); but in all probability the most useful aspects of further work along these lines is to be sought in particular applications to ground forms.

That this Adjunction Theorem breaks down for forms higher than the quadratic is perhaps one of the most remarkable facts of mathematics. It makes one wonder what would have been the history of geometry and natural philosophy, had the cubic or higher form been a possible absolute on which to base our metrical results. For never in the age-long story of measurement, from the discoveries of Pythagoras, about 500 B.C., to present-day speculations, has the geometer or physicist renounced the quadratic as his basis of measurement. The quadratic is one of the things which seem to have come to stay. The theorem concerning the squares on the sides r, x_1, x_2 of a right-angled triangle, we can write as

---

[1] Weitzenböck, *Ency. Math. Wiss.*, III, **3**, **6** (1922), p. 20. Burchardt, *Math. Annalen*, **43** (1893), 197–215.

$$r^2 = x_1{}^2 + x_2{}^2 = (x \mid x);$$

but the latest speculations in general relativity would throw this theorem into its infinitesimal shape

$$ds^2 = dx_1{}^2 + dx_2{}^2 = (dx \mid dx)$$

as a special case of a universal formula

$$ds^2 = \Sigma g_{ik}\, dx_i\, dx_k = g_{dx}{}^2$$

in n variables $x_1, x_2, \ldots, x_n$.

And what again would have happened if the absolute had been not even quadratic but only linear?

Why it is that the quadratic form should occupy this privileged position between linear and higher orders might well raise questions of considerable philosophic interest.

12. Connexion between Differential and Projective Invariants.

It may be wondered why there has hitherto been such preoccupation with the *linear* transformation, which after all is only a very special case of what in general can be written

$$T: \quad \left.\begin{array}{c} x_i = f(x_1', x_2', \ldots, x_n') = x_i(x_1', x_2', \ldots, x_n') \\ i = 1, 2, \ldots, n \end{array}\right\}, \qquad (34)$$

where a set of independent variables x_i is transformed into a new set $x_1', \ldots, x_n'$ by definite functional relations not necessarily linear. Now the reason lies in the general difficulty of treating anything more elaborate. The linear stands in relation to the general transformation T, much as linear differential equations do to the general theory of differential equations, or in kinematics as the velocity of a particle to a finite displacement. The latter may disclose quite an unworkable problem to which the former contributes a satisfactory first approximation.

Assuming each x_i to be a regular function of each x_j', and vice versa, we can write

$$dx_i = \frac{\partial x_i}{\partial x_1'}\, dx_1' + \ldots + \frac{\partial x_i}{\partial x_n'}\, dx_n'$$

for the first differential of each x_i in terms of those of x_i'. But *this is patently a linear transformation from the set*

$$\{ dx_1,\ dx_2,\ \ldots,\ dx_n \}$$

to $\{dx_1', dx_2', \ldots, dx_n'\}$. Let us denote these sets by dx, dx' respectively. Then if $c_1, \ldots, c_n$ are given functions of $x_1, \ldots, x_n$, we can express each c_i as a function of $x_1', \ldots, x_n'$, and any linear form $\sum\limits_{i=1}^{n} c_i dx_i$ as a linear form in dx_i'. We write

$$C = \sum_{i=1}^{n} c_i dx_i = (c \mid dx) = (c' \mid dx'),$$

where a new set of functions c_i' is derived as coefficients of dx_i' in C. But this is precisely the theory of contragredience over again, and we can accordingly speak of the set of functions

$$c = [c_1, c_2, \ldots, c_n]$$

as contragredient to the set dx.

Example.—

Writing in matrix notation $dx = M dx'$, $c = c'M^{-1}$, then

$$M = \left[\frac{\partial x_i}{\partial x_j'}\right] = \left[\frac{\partial x}{\partial x'}\right], \quad M^{-1} = \left[\frac{\partial x_i'}{\partial x_j}\right], \quad MM^{-1} = I,$$

where $|M| = \dfrac{\partial(x)}{\partial(x')} \neq 0$, $|M|^{-1} = \dfrac{\partial(x')}{\partial(x)} \neq 0$ (cf. VI, p. 126).

Thus:

Arising out of a **general** *transformation* $\mathrm{x} \longrightarrow \mathrm{x}'$ *is a* **linear** *transformation* $\mathrm{dx} \longrightarrow \mathrm{dx}'$ *whose matrix* $\mathrm{M} = \left[\dfrac{\partial \mathrm{x_i}}{\partial \mathrm{x'_j}}\right]$ *is non-singular, together with a contragredient transformation* $\mathrm{c} \longrightarrow \mathrm{c}'$ *for the set* c *of coefficients of a linear differential form* $\Sigma \mathrm{c_i dx_i}$.

In particular let $f = f(x) = f'(x')$ denote a given function expressed first explicitly in terms of the x_i's and secondly in terms of the x_i''s. Then its total differential can be written as

$$df = \Sigma \frac{\partial f}{\partial x_i} dx_i = \Sigma \frac{\partial f'}{\partial x_i'} dx_i' = df',$$

or

$$\left(\frac{\partial f}{\partial x}\middle| dx\right) = \left(\frac{\partial f'}{\partial x'}\middle| dx'\right).$$

So, whatever function f is taken, the set

$$\left\{\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right\}$$

is contragredient to $[dx_1, dx_2, \ldots, dx_n]$.

Once this algebraic idea is grasped—that an inner product $\left(\frac{\partial f}{\partial x}\, dx\right)$ is an invariant for contragredient sets, algebraic or differential—it throws light on numerous branches of geometry and physics, bringing them under the rubric of one mathematical doctrine. Thus it appears that

Binary forms illustrate the differential geometry on a surface.

Ternary forms illustrate that of pre-relativity physics.

Quaternary forms illustrate the present era of physics.

EXAMPLES

1. The well-known formula

$$ds^2 = dx^2 + dy^2$$

for the square of the element of arc of a plane curve in terms of differentials in rectangular co-ordinates x, y, can be looked upon as a binary quadratic ground form in homogeneous variables $[dx, dy]$.

2. The analogous ternary formula

$$ds^2 = dx_1{}^2 + dx_2{}^2 + dx_3{}^2$$

for the arc of a space curve is again a quadratic. Then if a linear transformation $dx \longrightarrow dx'$ leaves this absolutely invariant, we have another example of orthogonal transformation. In this case

$$ds^2 = (dx \mid dx) = (dx' \mid dx') = ds'^2, \qquad dx = M dx',$$

and M is orthogonal.

3. Or, again, the potential function V of three variables x_1, x_2, x_3 leads to the important vector $\left\{\frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \frac{\partial V}{\partial x_3}\right\}$ contragredient to (dx_1, dx_2, dx_3).

A full account of this differential theory can be read in many recent publications.[1] But here it may be useful to refer to the likenesses and the contrasts between the algebraic and the differential theories. Both contain ground forms, linear transformations, and concomitants relative or absolute, though they are somewhat disguised by having different names. Certain processes, too, can be recognized as identical. But in the differential theory emphasis is placed on the *tensor* (pp. 91, 200), or *set* of coefficients $a_{ijk\ldots}^{rst\ldots}$ of a multilinear form where $ijk\ldots$ are called *indices of covariance* and $rst\ldots$ *indices of contravariance.* Thus a set a_i is called a *covariant vector* and a^r, a *contravariant vector.*

---

[1] An excellent introduction is given in the Cambridge Tract, No. 24 (1927), Veblen, *Invariants of Quadratic Differential Forms*. A larger work is the English translation of an Italian work: Levi-Civita, *The Absolute Differential Calculus* (Blackie, 1926). Cf. Weitzenböck, *Invariantentheorie*, Chap. XIII.

It should be carefully noted that this use of the words covariant and contravariant is quite different from their use in algebra.

From the algebraic point of view the most interesting fact of the differential theory of forms is the existence of a Reduction Theorem first discovered by Christoffel,[1] whereby the problem in differential invariant theory of tensors and their derivatives up to a given order, is identical with that of the projective invariant theory. It is noteworthy that from a physical point of view the most important algebraic forms are the linear a_x, u_α, the linear complex $(ab \mid xy)$, the quadratic $g_x{}^2 = \Sigma g_{ij} x_i x_j$, and the quadratic complex $(B \mid xy)^2 = \Sigma B_{ij,\, kl}\,(xy)_{ij}(xy)_{kl}$. These quadratics figure prominently as the differential form $ds^2 = \Sigma g_{ij} dx_i dx_j$ and the Riemann-Christoffel curvature tensor $B_{ij,\, kl}$.

13. Prepared Systems.

Although the general theory of binary forms is fairly complete, little is known of higher categories beyond the irreducible system of a ternary cubic and certain linear or quadric systems. The fundamental theorem works very well for ternary forms, because (abc), a_α, $(\alpha\beta\gamma)$ are the only types of symbolic factor which may arise even if compound co-ordinates are utilized (§11, p. 210). Quaternary forms require implicit convolution (§11, p. 211) and thereby provide great complications. How, for instance, do the symbols of a linear complex $(aa' \mid xy)$ fit in with the types $(abcd)$, a_α, $(\alpha\beta\gamma\delta)$? This has suggested the problem *to supplement the three fundamental symbolic types by further types so as to render all convolutions explicit.* To such systems of symbolic types the name Prepared Systems has been given.

---

[1] *Crelle,* **70** (1869), 46–70.

The prepared system for quaternary forms[1] consists of thirteen types:

$$(abcd),\ (\alpha\beta),\ (\alpha\beta\gamma\delta),\ (aAb),\ (\alpha A\beta),\ (AB),\ (aAB\alpha),$$
$$(aABCb),\ (\alpha ABC\beta),\ (aABCD\alpha),$$
$$(aABCDEb),\ (\alpha ABCDE\beta),\ \text{and}\ (ABCDEF).$$

Here capital letters have currency two (§6, p. 37); and if $A = a'a''$, $B = b'b''$, &c., then $(aABCb)$ is defined as

$$(aAb')\,(b''Cb) - (aAb'')\,(b'Cb)$$

with analogous definitions for the others. A prepared system in general gives the complete system for all possible *linear* ground forms $a_x, \ldots, u_\alpha, \ldots, (A_2\pi_2), \ldots, (A_r \mid \pi_r), \ldots$, but at present nothing is known beyond the quaternary case.

14. Quantitative Substitutional Analysis.

Determinant, matrix, symbolic invariant, tensor, and group theory are but variations on one theme—permutations and combinations. Here algebra begins and here it appears to stay. But how many of us have ever thought it worth while to study the very ABC of substitutional processes? or have even inquired if they *have* an ABC?

Suppose, for instance, it is known that a certain function $f(x, y, z, u, v)$ of five arguments is symmetrical in x, y, skew symmetrical in y, z, u and also in x, v: then what are its characteristic properties? Is there a calculus behind this kind of inquiry which will obviate the necessity of examining every case for itself as it arises? There is. About thirty years ago Frobenius[2] and Young[3] appear to have made independent discoveries which lead to a systematic calculus of substitutions. Their work links these questions with the theory of matrix equations. By so doing, it gives a kind of canonical form to whole groups of substitutional properties. The bare fact that the natural sequence in this algebra is not of the order 1, 2, 3, ... but rather is that of 1!, 2!, 3!, . . . shows where the practical difficulties lie. Napier was prompted to invent logarithms solely by the difficulty of computing long multiplication sums. Can a like benefit, it may be asked, be found for algebra, and have these pioneers brought it within sight?

---

[1] *Proc. London Math. Soc.* (2), **21** (1923), 381–88, and (2), **25** (1926), 303–327.

[2] "Über die Darstellung der endlichen Gruppen durch lineare Substitutionen", *Berliner Sitzungsberichte*, **1** (1897), **2** (1899).

[3] *Proc. London Math. Soc.* (1), **33** (1901), **34** (1903), (2), **28** (1928), **31** (1930), **34** (1932), **35** (1933), **36**, **37** (1934), and *Journal*, **3** (1928), "On Quantitative Substitutional Analysis". Cf. *Phil. Trans. A.*, **234** (1935).

MISCELLANEOUS EXAMPLES

1. If capital letters denote square matrices of order n, I being the unit matrix, and if small letters denote scalar numbers, prove that

$$(I + pAB)A(I + qBA) = (I + qAB)A(I + pBA).$$

2. Prove that $(I + pAB)A(I + qBA)B(I + rAB)$ is symmetrical in p, q, r.

3. If $A : B$ means AB^{-1}, prove $A : B = AX : BX$, if $|BX| \neq 0$.

4. If $\dfrac{A}{B+}\,\dfrac{C}{D+}\dots$ has the usual meaning, of a continued fraction, provided that division is always performed on the right, investigate the law of successive convergents, P/Q.

[With proper safeguards, the usual scalar law is true. If $A = C = \dots = I$, $B = A_1$, $D = A_2$, . . . , then $P_n = P_{n-1}A_n + P_{n-2}$, $Q_n = Q_{n-1}A_n + Q_{n-2}$.

5. Given

$$A = \begin{bmatrix} 4 & -1 & 0 \\ -1 & 3 & 2 \\ 5 & 7 & 6 \end{bmatrix} \text{ and } B = \begin{matrix} -1 & 12 & 5 \\ 12 & 11 & 9 \\ 34 & 57 & 37 \end{matrix}$$

find the most general matrix X, satisfying the equation $AX = B$

(*Edinburgh.*)

6. Show that $A = \begin{bmatrix} 0 & 1 & 2 & 3 \\ -1 & 0 & 4 & 7 \\ -2 & -4 & 0 & 2 \\ -3 & -7 & -2 & 0 \end{bmatrix}$ satisfies the equation

$A^3 + 83A = 0$. Obtain the Cayley-Hamiltonian equation (§2, p. 99) for A, and discuss the connexion between the two equations.

(*Edinburgh.*)

7. If θ, φ, ψ are arbitrary then the matrix

$$\begin{bmatrix} \cos\varphi\cos\theta\cos\psi - \sin\varphi\sin\psi, & \sin\varphi\cos\theta\cos\psi + \cos\varphi\sin\psi, & -\sin\theta\cos\psi \\ -\cos\varphi\cos\theta\sin\psi - \sin\varphi\cos\psi, & -\sin\varphi\cos\theta\sin\psi + \cos\varphi\cos\psi, & \sin\theta\sin\psi \\ \cos\varphi\sin\theta, & \sin\varphi\sin\theta, & \cos\theta \end{bmatrix}$$

is orthogonal. Express it in the form $(I - S)(I + S)^{-1}$, where S is skew symmetric.

(*Edinburgh.*)

8. Prove that

$$A = \begin{bmatrix} 1 & 1 & . \\ . & 1 & . \\ . & . & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & -1 & . \\ . & 1 & . \\ . & . & 1 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & . & . \\ . & 2 & . \\ . & . & 1 \end{bmatrix}$$

are commutative, and that the functions $A + B + C$, $BC + CA + AB$, ABC are scalar, and equal to the corresponding functions of the latent roots of A. (*Edinburgh.*)

9. Prove that if $A = \begin{bmatrix} 1 & -1 & 1 & -1 & 1 \\ 4 & -3 & 2 & -1 & \\ 6 & -3 & 1 & . & \\ 4 & -1 & . & . & \\ 1 & . & . & . & \end{bmatrix}$ then $A^3 = I$, and generalize the theorem. Prove that the characteristic equation satisfied by A is $(z^3 - 1)^2 (z - 1)^{-1} = 0$.

10. If n, m, k are positive integers, and a function ω satisfies the relation

$$\omega(n, m) + \omega(m, k) = \omega(n, k),$$

show that $\omega(n, m)$ is the (n, m)th element in a skew symmetric matrix.

[illegible].)

11. If $A_\alpha = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$ prove $A_\alpha{}^n = \begin{bmatrix} \cos n\alpha, & \sin n\alpha \\ -\sin n\alpha, & \cos n\alpha \end{bmatrix}$ and give a geometrical explanation.

If A_β is a similar matrix for an angle β, prove A, B commute and that $A_\alpha A_\beta = A_{\alpha+\beta}$.

12. If $A = \begin{bmatrix} a & b & c \\ . & a & b \\ . & . & a \end{bmatrix}$, $P = \begin{bmatrix} p & q & r \\ . & p & q \\ . & . & p \end{bmatrix}$ then AP, PA both have the same form, with constant values throughout the parallels to the leading diagonal, and zeros below.

Generalize this feature.

13. If A, B are reduced to normal form PLP^{-1}, QMQ^{-1} where L and M are diagonal matrices both with n distinct latent roots $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $\mu_1, \mu_2, \ldots, \mu_n$, prove that a non-zero matrix X can be found to satisfy

$$AX = XB$$

if and only if A and B have at least one latent root in common. In this case X is called a commutant of A and B.

[Let $Y = P^{-1}XQ$; then $LY = YM$; and equate corresponding elements.

14. If $AX = XB$, prove that X is a commutant of $f(A)$ and $f(B)$.

[First prove $A^nX = XB^n$.

15. *Desargues' Theorem.*—In ternary forms, let a, b, c, a', b', c' denote coplanar lines. Show that the equation

$$p_x = (b'c'b)c_x - (b'c'c)b_x = 0$$

represents the line joining the intersections of b, c and of b', c'. If q_x, r_x denote similar expressions for c, a and a, b prove that

$$(pqr) = -(abc)(a'b'c')(aa'\,.\,bb'\,.\,cc')$$

where

$$(aa'\,.\,bb'\,.\,cc') = (abb')(a'cc') - (a'bb')(acc').$$

[Write $p = \lambda c - \mu b$, $q = \lambda' a - \mu' c$, $r = \lambda'' b - \mu'' a$, where $\lambda = (b'c'b)$, &c., are scalar. Then $(pqr) = \lambda\lambda'\lambda''(cab) - \mu\mu'\mu''(bca)$, all other terms vanishing. The result follows by using the identities

$$(b'c'b)(c'a'c) = (b'c'a')(c'bc) + (b'c'c)(c'a'b)$$
$$(c'a'a)(a'b'b) = (c'a'b')(a'ab) + (c'a'b)(a'b'a).$$

State the geometrical result, and the dual result involving points α, β, γ, α', β', γ'.]

16. Prove the identity

$$(ax\,.\,by\,.\,cz) + (\text{[illegible]}\,.\,bz\,.\,cx) + (az\,.\,bx\,.\,cy) = 0$$

between any six coplanar *points* a, b, c, x, y, z. The symbols are cogredient and equally well can represent lines.

Pappus' Theorem.—For six coplanar points A, B, C, X, Y, Z, if AX, BY, CZ meet in a point, and AY, BZ, CX meet in a point, prove that AZ, BX, CY do so.

[Expand each compound determinant.

17. If $(\xi ax\,.\,by\,.\,cz\xi) \equiv (\xi axb)(ycz\xi) - (\xi axy)(bcz\xi)$ where the symbols are seven cogredient points in the quaternary field, prove that this expression equated to zero is the equation in ξ of the quadric surface containing the lines ax, by, cz as generators.

[The point $\lambda a + \mu x$ lies on the surface, &c.

18. Prove the identity

$$(\xi ax\,.\,by\,.\,cz\xi) + (\xi ay\,.\,bz\,.\,cx\xi) + (\xi az\,.\,bx\,.\,cy\xi) = 0.$$

If $ABCXYZ$ is a given skew hexagon in space, prove that the quadric surface containing AX, BY, CZ as generators, and that containing AY, BZ, CX as generators, and that containing AZ, BX, CY as generators are linearly related and have a common curve of intersection.

CHAPTER XXII

MISCELLANEOUS PROPERTIES OF DETERMINANTS

1. Remark on the Fundamental Identities.

In the first edition of this book the formula

$$(abc\dot{d})\,(\dot{e}\dot{f}gh) = (\dot{a}def)\,(\dot{b}\dot{c}gh) \quad . \quad . \quad . \quad . \quad (1)$$

appeared at (30) and (31) of p. 43, but with a wrong sign throughout on the right-hand side. This serves to illustrate the importance of stating a ready rule for fixing the signs in any such identity. Such a rule is this: the arrangement of *all* the single letters $a, \ldots, h$ of every term on the left must belong to one class, say C_+, while that of every term on the right must belong to the opposite class, say C_-. If so, no further sign is needed. The proof which leads to identities (33), (34) of p. 43 and I, II, III of p. 45, and so on, really secures this, and it ultimately depends on the $2n$-rowed determinants which involve all the letters of the pair of n-rowed determinants occurring in a term of the series.

In practice, on being given the series on the left in the above identity, we may obtain that on the right with its correct signs by interchanging e with b and f with c, and then removing all the dots from def and placing dots over *all* the once undotted letters abc of the first factor. Again, each of the above series is equal to a further series

$$(abc\dot{g})\,(efd\dot{h}) \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \quad (2)$$

which is written down by assembling d along with ef in the second factor, thus displacing g, and then transferring the dots from def to gh.

A good illustration of these operations is to regard the matrices $A = abc$, $B = def$, $C = gh$ as three sets of library volumes, all of equal thickness, arranged to fill two equal shelves. The obvious fact that one such set must be broken, and that this may be either A or B or C, is a helpful clue to the existence of the three modes of the same expression as given in (1) and (2), if it be noted that the constituents of the broken set receive the dots. Explicit convolution (p. 46) is the assembling of a complete set upon one shelf, while implicit convolution indicates that a set is distributed over two or more shelves.

Each identity is the result of assembling a given broken set upon any one shelf. For instance, if a third factor $(pqrs)$ occurs, we may assemble B in this factor and obtain

$$(abc\dot{d})\,(\dot{e}\dot{f}gh)\,(pqrs) = (abc\dot{p})\,(\dot{q}\dot{r}gh)\,(def\dot{s}).$$

This follows from a twelve-rowed determinant Δ_3 expressed in two equivalent ways as in (40), p. 47.

2. Partitions by Rows and Columns. Double Currency.

At p. 37 currency has been defined by columns. It is useful to extend the definition to rows also. Just as A and B have been used to denote the first r and the next $n - r$ columns of a determinant (or matrix) of n columns, so let X and Y denote the first s and next $n - s$ rows. More precisely let the determinant be written

$$\Delta = \begin{vmatrix} A_X & B_X \\ A_Y & B_Y \end{vmatrix}$$

where, for example, $A = ab$, $B = cd$, $X = 123$, $Y = 4$, and

$$| a_1 b_2 c_3 d_4 | = \begin{matrix} ab_{123} & cd_{123} \\ ab_4 & cd_4 \end{matrix}$$

Here ab_{123} denotes the submatrix consisting of the first two columns and first three rows of the whole determinant (or matrix), with analogous meanings for the other submatrices. Evidently this process of partitioning by rows can proceed independently of that by columns.

3. Conformable Partitions for Addition and for Multiplication.

It is an immediate consequence of the definition of matrix addition that, when two matrices are partitioned into submatrices in the same way by rows, and also in the same way by columns, then the same law of addition applies to the submatrices as to the elements composing them. Two such matrices are said to be *partitioned conformably for addition.*

For example, $[P, Q] + [R, S] = [P + R, Q + S]$ provided that the submatrices P and R have the same number of columns, and so have Q and S, while all have the same number of rows.

Two matrices are *partitioned conformably for multiplication* when the column partitions of the fore-factor agree with the row partitions of the after-factor. This at once allows the rule of multiplication by row into column to operate without further breaking the partitions.

For example $[P, Q]\begin{bmatrix} V \\ W \end{bmatrix} = [PV + QW]$ provided that P has the same number of columns as V has rows, and Q has the same number of columns as W has rows. Also P, Q have the same number of rows and V, W the same number of columns. This follows at once from the definition of matrix multiplication. More complicated examples of partitioning both by rows and by columns, whether for addition or for multiplication, are easy to construct.

4. Schweins' Theorem.

Consider a determinant from which a rectangular (or square) block of elements have been removed, as for example

$$\begin{array}{cccccc} a_1 & b_1 & c_1 & \cdot & \cdot & \cdot \\ a_2 & b_2 & c_2 & \cdot & \cdot & \cdot \\ a_3 & b_3 & c_3 & d_3 & e_3 & f_3 \\ a_4 & b_4 & c_4 & d_4 & e_4 & f_4 \\ a_5 & b_5 & c_5 & d_5 & e_5 & f_5 \\ a_6 & b_6 & c_6 & d_6 & e_6 & f_6 \end{array} \qquad (3)$$

If this is expanded by the $\left(\frac{2}{4}\right)$ Laplace development it yields $\Sigma(ab)_{12}(cdef)_{3456}$: and since neither suffix 1 nor 2 can occur with d, e or f, we infer that a, b, c alone are deranged, and

$$\Delta = (\dot{a}\dot{b})_{12}(\dot{c}def)_{3456}.$$

By the same reasoning the $(3 \,\vdots\, 3)$ development yields

$$\Delta = (abc)_{12\dot{3}}(def)_{\dot{4}\dot{5}\dot{6}}.$$

The consequent identity

$$(\dot a\dot b)_{12}(\dot c def)_{3456} = (abc)_{12\dot 3}(def)_{\dot 4\dot 5\dot 6},$$

which is obviously akin to the Sylvester identities of an early chapter, was discovered by Schweins as early as the year 1825 in all its generality. We take such a determinant Δ of n rows and columns, with a vacant array in the first p rows and the last r columns. The corresponding Laplace development then yields the identity

$$\Delta = \begin{array}{ccc} A_X & B_X & \\ A_Y & B_Y & C_Y \\ A_Z & B_Z & C_Z \end{array} = A_X(BC)_{YZ} = (AB)_{X\dot Y}C_{\dot Z}. \quad (4)$$

This is Schweins' Theorem, although his original statement and proof were much more cumbersome. Here A_X, B_Y, C_Z are square arrays of p, q, r rows and columns respectively, so that

$$p + q + r = n.$$

In the above example $n = 6$, $p = 2$, $q = 1$, $r = 3$, $A = ab$, $B = c$, $C = def$, $X = 12$, $Y = 3$, $Z = 456$. If $q = 0$ then Δ merely factorizes into (A_X) (B_Y).

EXAMPLES

1. Expand $\begin{array}{cccc} & b_1 & c_1 & \\ a_2 & & c_2 & \\ a_3 & b_3 & c_3 & d_3 \end{array}$ and deduce that $(abc)_{12\dot 3}d_{\dot 4} = (\dot a\dot b)_{12}(\dot c d)_{34}$.

2. $(abcd)_{\dot 1\dot 2\dot 3\dot 4}(efg)_{\dot 5 67} = (abcd\bar e)_{12345}(\bar f\bar g)_{67}$.

[Place the zeros in the lower left block of Δ.

3. $(abcd)_{\dot 1\dot 2\dot 3\dot 4}(efa)_{\dot 5 67} = (abcd\dot e)_{12345}(\dot f a)_{67}$.

4. $(abcd)_{1\dot 2\dot 3\dot 4}(eab)_{\dot 5 61} = (abcde)_{12345}(ab)_{61}$.

All these identities have arisen through equating two separate Laplace developments of Δ, one by rows and one by columns. Had these operations proceeded in succession, still another form of the expansion would emerge. For instance in Ex. 1 we should then obtain $(\dot a\dot b)_{12}\dot c_{\bar 3}d_{\bar 4}$, where separate marks $\cdot$ and $-$ must now be used, to denote the separate determinantal permutations first of ab, c and next of 3, 4. In all it is a series of six terms. In general

$$\Delta = \dot{A}_X \dot{B}_{\bar{Y}} C_{\bar{Z}}.$$

We shall call this a *double Laplace development* of the determinant. It applies usefully whenever Δ has a square or rectangular block of zeros: and the method was used on p. 102 for a bordered determinant, with a square block of zeros.

5. Further Use of the Unit Matrix.

Let $(123 \ldots n)$ denote the determinant $| \delta_{ij} |$ of the unit matrix, where 1 denotes the column $\{1, 0, \ldots, 0\}$, 2 denotes $\{0, 1, 0, \ldots, 0\}$ and so on. This gives a very easy way of expressing the cofactor for any element in a general determinant. Thus, in $(abcd)$,

$$B_3 = (a3cd) = \begin{matrix} a_1 & \cdot & c_1 & d_1 \\ a_2 & \cdot & c_2 & d_2 \\ a_3 & 1 & c_3 & d_3 \\ a_4 & \cdot & c_4 & d_4 \end{matrix}$$

for on multiplying the second column here by b_3 we obtain the expression which is equal to Δ when $b_1 = b_2 = b_4 = 0$. This gives a useful rule for denoting the cofactor of m_i in Δ: namely, replace m_i of column m by the integer i.

EXAMPLES

1. What cofactors do $(1bc)$, $(2bc)$, $(a3c)$, $(ab1)$ represent in the three-rowed determinant $(abc) = | a_1 b_2 c_3 |$? Write them out explicitly.

Ans. Those of a_1, a_2, b_3, c_1.

2. Verify that $(12cd)$ represents the cofactor of $(ab)_{12}$ in $(abcd)$. Generalize this feature.

3.
$$\begin{vmatrix} A_1 & B_1 \\ A_2 & B_2 \end{vmatrix} \quad (\bar{1}bcd)\,(a\bar{2}cd).$$

$$\begin{vmatrix} A_1 & C_1 & D_1 \\ A_2 & C_2 & D_2 \\ A_4 & C_4 & D_4 \end{vmatrix} = (\bar{1}bcd)\,(ab\bar{2}d)\,(abc\bar{4}).$$

The above examples throw certain minors occurring in the adjugate determinant $| A_1 B_2 C_3 D_4 |$ into a new form, let us say the dilated form, where every minor is expressed as a determinant of the same order as the original. By the fundamental identities we obtain, from Ex. 3 and 4,

$$(12cd)\ (abcd) \text{ and } (1b24)\ (abcd)^2$$

respectively. But

$$(12cd) = (cd)_{34}, \quad (1b24) = \left|\begin{array}{cccc} 1 & b_1 & & \\ & b_2 & 1 & \\ & & & \\ & b_4 & . & \end{array}\right.$$

Hence $\quad | A_1 B_2 | = | c_3 d_4 | \Delta, \quad | A_1 C_2 D_4 | = - b_3 \,.\, \Delta^2. \qquad (5)$

This method gives another proof of Jacobi's theorem on the adjugate (p. 79), and it also indicates how to prove the general theorems epitomized in the examples of p. 56.

6. Schweins' Series.

Besides the theorem already given, Schweins found an important series, in 1825, for the expansion of the quotient of two n-rowed determinants which differ only in one column, say of $(ab \ldots cd)/(ab \ldots ce)$. This series is of great use in many branches of algebra and analysis; and many interesting cases arise by treating d or e as a column of the unit matrix. The series, which depends upon the repeated application of Jacobi's theorem on the adjugate, is illustrated by the following examples.

EXAMPLES

1. By applying Schweins' Theorem to the determinant

$$\begin{array}{cccccc} a_1 & b_1 & c_1 & & & \\ a_2 & b_2 & c_2 & & & \\ a_1 & b_1 & c_1 & a_1 & b_1 & d_1 \\ a_2 & b_2 & c_2 & a_2 & b_2 & d_2 \\ a_3 & b_3 & c_3 & a_3 & b_3 & d_3 \\ a_4 & b_4 & c_4 & a_4 & b_4 & d_4 \end{array}$$

prove that $(abc)_{123}(abd)_{124} = (ab)_{12} \times (abcd)_{1234}$, which is a particular case of Jacobi's theorem on the adjugate.

[Use the $\frac{2}{4}$ and $3 \,\vdots\, 3$ developments.

2. Provided that none of b_2, $(bc)_{23}$, $(bcd)_{234}$ vanish, prove that

$$\frac{(abcd)_{1234}}{(bcd)_{234}} = a_1 - \frac{a_2 \,.\, b_1}{1 \,.\, b_2} - \frac{(ab)_{23}(bc)_{12}}{b_2(bc)_{23}} - \frac{(abc)_{234}(bcd)_{123}}{(bc)_{23}(bcd)_{234}},$$

which is a case of *Schweins' Series.*

[The sums of the first two, three, four terms, respectively, of this series are $\dfrac{(ab)_{12}}{b_2}$, $\dfrac{(abc)_{123}}{(bc)_{23}}$, $\dfrac{(abcd)_{1234}}{(bcd)_{234}}$, by Jacobi's theorem.

3. $$\frac{(abcd)_{1234}}{(abce)_{1234}} = \frac{(abc)_{123}(abed)_{1234}}{(abe)_{123}(abce)_{1234}} + \frac{(ab)_{12}(aed)_{123}}{(ae)_{12}(abe)_{123}} + \frac{a_1(ed)_{12}}{e_1(ae)_{12}} + \frac{d_1}{e_1}.$$

(*Schweins' Series*, 1825.)

4. $$\frac{(abd)}{(abc)} = \frac{(abx)(acd)}{(acx)(abc)} + \frac{(axy)(cdx)}{(cxy)(acx)} + \frac{(dxy)}{(cxy)},$$

$$\frac{(abc)}{(abz)} + \frac{(ayb)(acz)}{(abz)(ayz)} + \frac{(axy)(cyz)}{(ayz)(xyz)} = \frac{(xyc)}{(xyz)}.$$

[Use the fundamental identities. On putting the unit matrix for xyz these series revert to series of types shown in Ex. 2 and 3.

5. Generalize the above results.

[In Ex. 1, take $m \times (m+1)$ for the size of the zero block. In Ex. 2 the left-hand denominator is an n — 1-rowed minor of the n-rowed numerator.

7. Dodgson's Method of Evaluating a Determinant by Condensation.

Form the sequence of determinants whose orders are indicated by their suffixes:

$$\Delta_n = \begin{array}{lll} a_1 & b_1 & \ldots \\ a_2 & b_2 & \ldots \end{array}, \quad \Delta_{n-1} = \begin{array}{lll} (ab)_{12} & (bc)_{12} & \ldots \\ (ab)_{23} & (bc)_{23} & \ldots \\ \cdot & \cdot & \cdot \end{array}$$

$$\Delta_{n-2} = \begin{array}{ll} (abc)_{123} & (bcd)_{123} \\ (abc)_{234} & (bcd)_{234} \end{array}$$

then the last Δ_1 is a single term $(ab\ldots)_{12\ldots}$ which is Δ_n. These are formed by cross multiplication of each set of four *adjacent* elements and then dividing by their common *pivotal* element. It is the technique already used in Ex. 2 above. Thus, in Δ_{n-2}, $(abc)_{123}$ arises from the leading two-rowed minor of Δ_{n-1} divided by b_2. Aitken modified the method to suit the case when it fails through the vanishing of this pivot.

Example.—

Evaluate the four-rowed alternant $a^0b^1c^2d^3\,|$ directly by Dodgson's method.

8. The Bordered Unit Matrix and the Binet-Cauchy Theorem.

The bordered unit matrix, which is an important type of bordered determinant, may be written

$$\Delta = \begin{array}{cc} & A \\ X & I \end{array} = (-)^r \Sigma (ab\ldots)_{ij\ldots}(xy\ldots)_{ij\ldots}, \qquad (6)$$

where I denotes the $n \times n$ unit matrix, X is a set of r columns $x, y, \ldots$ each of n elements, and A is a set of r rows $a, b, \ldots$ each of n elements. The proof follows at once by a double Laplace development as before; and the summation has $\binom{n}{r}$ terms, due to the combinations of r distinct suffixes in a term.

Example.—

$$\begin{array}{cccccc} \cdot & \cdot & a_1 & a_2 & a_3 & a_4 \\ \cdot & \cdot & b_1 & b_2 & b_3 & b_4 \\ x_1 & y_1 & 1 & \cdot & \cdot & \cdot \\ x_2 & y_2 & \cdot & 1 & \cdot & \cdot \\ x_3 & y_3 & \cdot & \cdot & 1 & \cdot \\ x_4 & y_4 & \cdot & \cdot & \cdot & 1 \end{array} \qquad \underset{ij}{\Sigma}(ab)_{ij}(xy)_{ij}, \text{ a series of six terms.} \quad . \quad . \quad (7)$$

If $r = n$, $\Delta = (-)^n \mid A \mid \mid X \mid$, which is now a product of two n-rowed determinants; and if $r > n$ then Δ vanishes. We may, if we prefer, change the sign of each element in X, writing $-X$ instead; and this avoids the awkward sign preceding the summation. Thus, for odd or even r,

$$\begin{array}{cc} \cdot & A \\ -X & I \end{array} = \Sigma (ab\ldots)_{ij\ldots}(xy\ldots)_{ij\ldots} = (A \mid X). \qquad (8)$$

Hence the bordered unit determinant is expressible as an rth compound inner product, or bideterminant (p. 82), the width of the border being r.

Next, let I_r be the $r \times r$ unit matrix. Then, by matrix multiplication,

$$\begin{array}{cc} \cdot & A \\ -X & I \end{array} \quad \begin{array}{cc} I_r & \cdot \\ X & I \end{array} \qquad \begin{array}{cc} AX & A \\ -X+X & I \end{array} \qquad \begin{array}{cc} R & A \\ \cdot & I \end{array} \qquad (9)$$

where $R = AX$ is an $r \times r$ matrix. But $AX = \begin{bmatrix} a_x & a_y & \ldots \\ b_x & b_y & \ldots \\ \cdot & \cdot & \cdot \quad \cdot \end{bmatrix}$, while the two determinants in the left-hand side product (9) are $(A \mid X)$

and unity respectively; and the right-hand side of (9) is $| R |$ simply (p. 32). Hence

$$\begin{matrix} \cdot & A \\ -X & I \end{matrix} = | AX | = \begin{matrix} a_x & a_y \\ b_x & b_y \end{matrix} \tag{10}$$

By equating (8) and (10) we obtain the Binet-Cauchy theorem (p. 81), which can be written more succinctly as

$$(A \mid X) = | AX | \tag{11}$$

9. Bideterminants.

Let the bideterminant be written as follows:

$$(A \mid X) = \begin{array}{cccc|cccc} a_1 & b_1 & \dots & d_1 & x_1 & y_1 & \dots & z_1 \\ a_2 & b_2 & \dots & d_2 & x_2 & y_2 & \dots & z_2 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_n & b_n & \dots & d_n & x_n & y_n & \dots & z_n \end{array} \;. \tag{12}$$

with r columns on the left and r columns on the right of a central vertical barrier. Those on the left are the transposed of the matrix A; those on the right are the matrix X. This bideterminant therefore admits of an *expanded* form (8) in terms of r-rowed minors chosen in pairs from A and X, and also of a *determinantal* form

$$\Sigma \pm a_x b_y \dots d_z \tag{13}$$

as in (10). Several of its properties which have been given on p. 82 may now be supplemented: namely, *the bideterminant is unaltered by*

(i) the interchange of any pair of rows,
(ii) the operation $\text{row}_i + p\ \text{row}_j$, $\text{row}_j - p\ \text{row}_i$ applied to its left- and right-hand matrices respectively,
(iii) the insertion or removal of a completely zero row. (14)

The proof of (i) follows from (13), since all its parts $a_x (= \Sigma a_i x_i)$ are symmetrical in the row suffixes i. Next (ii) follows since

$$a_x = a_1 x_1 + \dots + (a_i + pa_j) x_i + \dots + a_j (x_j - px_i) + \dots + a_n x_n$$

for all values of p, while (iii) involves adding or subtracting zero terms only.

Bideterminants are useful in numerical cases, but we shall now consider one of their general properties which throws light on the dual form of the Sylvester identities.

Consider the identity

$$\begin{array}{cccccc|} a_1 & b_1 & c_1 & x_1 & y_1 & z_1 \\ a_2 & b_2 & c_2 & x_2 & y_2 & z_2 \\ a_3 & b_3 & c_3 & x_3 & y_3 & z_3 \\ \cdot & \cdot & c_1 & \cdot & \cdot & -z_1 \\ \cdot & \cdot & c_2 & \cdot & \cdot & -z_2 \\ \cdot & \cdot & c_3 & \cdot & \cdot & -z_3 \end{array} = \begin{array}{|ccc|ccc|} a_1 & b_1 & \cdot & x_1 & y_1 & z_1 \\ a_2 & b_2 & \cdot & x_2 & y_2 & z_2 \\ a_3 & b_3 & \cdot & x_3 & y_3 & z_3 \\ \cdot & \cdot & c_1 & x_1 & y_1 & \cdot \\ \cdot & \cdot & c_2 & x_2 & y_2 & \cdot \\ \cdot & \cdot & c_3 & x_3 & y_3 & \cdot \end{array} \qquad (15)$$

which follows from (ii) by the operation

upper — lower half matrix, lower + upper half matrix

applied to the bideterminant on the left. Both bideterminants can either be expanded or take determinantal form. The latter in each case gives

$$\begin{array}{ccc} a_x & a_y & a_z \\ b_x & b_y & b_z \end{array}, \text{ where } a_x = a_1x_1 + a_2x_2 + a_3x_3, \qquad (16)$$

which is a determinant of the type needed in Schweins' Theorem, with a zero block of elements (here a single element). The absence of c_z is due to cancelling, within the six-term series, of the relevant columns, as inspection will show. The former operation, expansion by three-rowed minors in pairs, yields the result

$$(abc)\,(xyz) - (ab\,|\,xy)\,(c\,|\,z) = (ab\,|\,yz)\,(c\,|\,x) - (ab\,|\,xz)\,(c\,|\,y), \qquad (17)$$

which the reader may verify by taking each bideterminant in the form $\Sigma\,(abc)_{ijk}\,(xyz)_{ijk}$, summed for all possible choices of three out of the six rows, and allowing for the zero elements. The identity just obtained is of course trivial; but it suggests a more general result. In fact, the same reasoning gives the identity

$$\begin{array}{cccccccc} a & b & c & d & x & y & z & t \\ \cdot & \cdot & c & d & & & -z & -t \end{array} \qquad \begin{array}{cccccccc} a & b & & & x & y & z & t \\ & & c & d & x & y & \cdot & \cdot \end{array} \qquad (18)$$

from the above operation, this time applied to an 8-columned and $2n$-rowed bideterminant, each letter denoting a column of n elements. Let each bideterminant be expanded by its four-rowed minors in pairs. On the left the terms due to all possible choices of four distinct rows among the upper n rows sum up to

$$(abcd \mid xyzt).$$

Three upper and one lower row $[\,.\;.\;\dot{c}_l d_l \mid .\;.\;-z_l\;-t_l]$ yield

$$-\,(abc \mid xy\bar{z})\,(\dot{d} \mid \bar{t}),$$

which is a series of four terms, where c, d are permuted independently of z, t, while the negative sign is due to the single column of z or t selected from the lower matrix. Two upper and two lower rows yield

$$+\,(ab \mid xy)\,(cd \mid zt)$$

only, while terms involving three or more of the lower rows vanish, owing to zeros in their first two columns. On the right only one term survives. Hence

$$\begin{aligned}(abcd \mid xyzt) - (ab\dot{c} \mid xy\bar{z})\,(\dot{d} \mid \bar{t}) + (ab \mid xy)\,(cd \mid zt)&\\ = (ab \mid zt)\,(cd \mid xy), \quad . \quad . \quad . \quad . \quad . \quad &(19)\end{aligned}$$

which proves identity (29) of p. 95 without recourse to its dual (27). The more general results (30) and (31) of the same page follow similarly from the identity

$$\left|\begin{array}{cc|cc} A_i & B_i & P_i & Q_i \\ \cdot & B_i & \cdot & -Q_i \end{array}\right| = \left|\begin{array}{cc|cc} A_i & \cdot & P_i & Q_i \\ \cdot & B_i & P_i & \end{array}\right. \qquad (20)$$

The identity

$$\left|\begin{array}{cccc|cccc} A & B & C & \cdot & X & Y & \cdot & \cdot \\ \cdot & B & C & D & \cdot & \cdot & Z & T \end{array}\right| = \left|\begin{array}{cccc|cccc} A & B & C & \cdot & & Y & Z & T \\ -A & \cdot & \cdot & D & & \cdot & Z & T \end{array}\right. \qquad (21)$$

which follows from the operation $\text{row}_2 - \text{row}_1 \mid \text{row}_1 + \text{row}_2$, leads by a similar expansion to the further result

$$(A\dot{B} \mid XY)\,(\dot{C}D \mid ZT) = \sum_i \epsilon(\dot{A}_i BC \mid XY\bar{Z}_j)\,(\dot{A}_k D \mid \bar{Z}_k \bar{T}). \qquad (22)$$

Here A and X each have p columns, B and Y have q, C and Z have r, and D and T have s columns, while all eight matrices have n rows. Moreover $A = A_i A_k$ when partitioned into its first i and next k $(= p - i)$ columns, while $Z = Z_j Z_k$, $j + k = q$. The summation runs through the values of i from p downwards (and of k from zero upwards), while $\epsilon = (-)^{k(q+r+1)}$. The series terminates as soon as either A or Z is exhausted, whichever happens first.

The most important case is that when C has one more column than A. For example, the identity

$$\begin{array}{cccccccccc} a & a' & b & c & c' & c'' & & x & x' & y \\ & & b & c & c' & c'' & d & & & \end{array}$$

$$\begin{array}{ccccccccccc} a & a' & b & c & c' & & x & x' & z & z' & z'' \\ -a & -a' & & & & d & & & z & z' & z'' \end{array}$$

on expansion takes the form

$$(aa'\dot{b} \mid xx'y)\ (\dot{c}\dot{c}'\dot{c}''d \mid zz'z''t)$$

$$= (aa'bcc'c'' \mid xx'y\dot{z}\dot{z}'\dot{z}'')\ (d \mid \dot{t}) - (\bar{a}bcc'c'' \mid xx'y\dot{z}\dot{z}')\ (\bar{a}'d \mid \dot{z}''\dot{t})$$

$$+ (bcc'c'' \mid xx'y\dot{z})\ (aa'd \mid \dot{z}'\dot{z}''\dot{t}). \quad . \quad . \quad . \qquad (23)$$

These identities are of course closely allied to those of p. 45. In each case the implicit convolution of B, C on the left is replaced by their explicit convolution on the right. The columns a of A are removed one at a time (but always appear convolved explicitly or implicitly), and a set of columns from Z and T is drawn into the right wing of the first factor to balance the currency of the left.

The determinantal form of each of the above bideterminants is readily found to have one or more characteristic blocks of zero elements.

Conversely any such determinant, with a rectangular block of zeros such as

$$\Delta = \begin{array}{cccccc} a_\xi & b_\xi & c_\xi & d_\xi & e_\xi & f_\xi \\ a_\eta & b_\eta & c_\eta & d_\eta & e_\eta & f_\eta \\ & & & d_\zeta & e_\zeta & f_\zeta \\ a_x & b_x & & d_x & \cdot & \cdot \\ a_y & b_y & & & & \\ a_z & b_z & c_z & & & \end{array} \qquad a_\xi = \ddot{\Sigma}\, a_i \xi_i, \ n \geqslant 6, \qquad (24)$$

can be written as a bideterminant

$$\Delta' = \begin{array}{ccccccccc} a & b & c & d & e & f & \xi & \eta & \zeta \\ \cdot & \cdot & \cdot & \cdot & e & f & & & \end{array} \quad -x \ -y \ -z \qquad (25)$$

as is easily verified. When Δ' is expanded systematically as in (16) above it gives

$$(abcdef \mid \xi\eta\zeta xyz) - (abcd\dot{e} \mid \xi\eta\zeta\bar{x}\bar{y})\,(\dot{f} \mid \bar{z}) \\ + (abcd \mid \xi\eta\zeta\bar{x})\,(ef \mid \bar{y}\bar{z}) = \Delta'' \text{ say.} \qquad (26)$$

Thus any determinant lacking a rectangular block of elements can be expanded in the form of a series (26). In each term both sets of letters $a \ldots f$ and $\xi \ldots z$ preserve their original order; but the sign changes at each successive term. In each term the letters *ef* are convolved, explicitly or implicitly, and so are *xyz*, corresponding to the original zero block in Δ. The series starts with all six letters of each set in one factor: thereafter one letter of each set is transferred to a second factor at each new term. The series ends as soon as one (or both) of the sets *ef*, *xyz* are completely transferred; in this case *ef*.

EXAMPLES

1.

$$\begin{array}{cccccc} a_1 & b_1 & c_1 & d_1 & e_1 & f_1 \\ a_2 & b_2 & c_2 & d_2 & e_2 & t_2 \\ a_3 & b_3 & c_3 & d_3 & e_3 & t_3 \\ a_4 & b_4 & c_4 & d_4 & & \\ a_5 & b_5 & c_5 & d_5 & & \\ a_6 & b_6 & c_6 & d_6 & & \end{array} = (abcdef)_{123456} - (abcd\dot{e})_{123\bar{4}\bar{5}} f_{\bar{5}} \\ + (abcd)_{123\bar{4}} (ef)_{\bar{5}\bar{6}}.$$

[Use the first six columns of the unit matrix for $\xi\eta\zeta xyz$.

2. Prove that

$$\begin{array}{cccc} a_1 & \underline{b}_1 & c_1 & \\ & & & \\ & \bar{b}_3 & \bar{c}_3 & d_3 \\ & b_4 & c_4 & d_4 \end{array} \qquad \left| \begin{array}{ccccccc} a & b & c & 1 & 2 & & \\ . & b & c & & & 3 & 4 \end{array} \right|$$

where the bideterminant has $2n$ rows ($n > 3$), and the numerals denote the first four columns of the $n \times n$ unit matrix.

3. Prove that each bideterminant of (21) can be written in the form

$$\left| \begin{array}{cccc} A_X & B_X & C_X & . \\ A_Y & B_Y & C_Y & . \\ . & B_Z & C_Z & D_Z \\ . & B_T & C_T & D_T \end{array} \right|$$

and give the corresponding determinants for (18), (20), (23) above.

10. Note on Matrices.

The formula $\lambda B - AB = f(\lambda)I$, which occurs on p. 100, shows that B can be written as $f(\lambda)I \div (\lambda I - A)$ provided that the divisor is a non-singular matrix, which is the case when λ differs from each latent root of A. Since this expression involves the matrices I and A only, it may be worked out in powers of λ as an ordinary long-division sum; and since B is a polynomial in λ without remainder, the expression $f(A)$ must, by the remainder theorem of elementary algebra, vanish. Each coefficient B_i in this polynomial is a polynomial in A, as is apparent either from the formulæ on p. 100 or from actual division. Thus B itself commutes with A.

In special cases the matrix A satisfies an equation $\psi(\lambda) = 0$ of degree less than its order n: for instance a scalar matrix $[\alpha\delta_{ij}]$ of order n always satisfies the *linear* equation $A - \alpha I = 0$. The equation of lowest order satisfied by a given matrix A is called its *reduced characteristic equation*, the theory of which is due to Frobenius.[1]

If λ is a latent root of an $n \times n$ matrix A then the matrix $A - \lambda I$ must be singular, for its determinant vanishes. Consequently its rows are linearly dependent and, let us say, $\Sigma x_i \operatorname{row}_i = 0$, where the coefficients x_i are not all zero. On writing x for the column of n elements x_i we have at once

$$(A - \lambda I)x = 0, \quad x \neq 0.$$

This gives $Ax = \lambda x$. Conversely if this relation holds and x is not identically zero, then λ must be a latent root of A: otherwise the matrix $B = A - \lambda I$ is non-singular so that B^{-1} exists, and $B^{-1}Bx = x$, which is impossible since $Bx = 0$, $x \neq 0$.

EXAMPLES

1. Prove that the latent roots of a real symmetric matrix must be real, and those of a real skew symmetric matrix must be pure imaginary or else zero.

Let $\bar{A}$ denote the result of changing the imaginary i to $-i$ wherever it occurs among the elements of the matrix A. We call $\bar{A}$ the complex conjugate of the matrix A. Likewise $\bar{x}$ is the complex conjugate of x. If $Ax = \lambda x$, then $\bar{A}\bar{x} = \bar{\lambda}\bar{x}$. Whence, by transposition, $\bar{x}'\bar{A}' = \bar{\lambda}\bar{x}'$. Thus, by post-multiplication with x,

$$\bar{x}'\bar{A}'x = \bar{\lambda}\bar{x}'x.$$

Now $\bar{x}'x = \Sigma\bar{x}_i x_i$, which is the sum of the squares of the moduli of the elements x_i. It is real, and can only vanish if all x_i vanish. Hence $x \neq 0$ implies $\bar{x}'x \neq 0$.

Again if A is real, $A = \bar{A}$, and if symmetric also, $A = \bar{A}'$. Thus $\bar{x}'Ax = \bar{\lambda}\bar{x}'x$. Also $\bar{x}'Ax = \lambda\bar{x}'x$, by pre-multiplication in the original equation.

---

[1] Cf. Turnbull and Aitken, *Canonical Matrices* (Glasgow, 1932).

Thus $\lambda\bar{x}'x = \bar{\lambda}\bar{x}'x$, so that $\lambda = \bar{\lambda}$ since the other factor is non-zero. Thus λ must be real.

Further if A is real and skew symmetric, $A = -\bar{A}'$, so that $\lambda = -\bar{\lambda}$, which means that λ is pure imaginary or else zero.

2. Deduce the result §8, p. 107.

3. If $A = \bar{A}'$, A is said to be Hermitian. Prove that its latent roots must be real.

4. If $A'A = I$, so that A is orthogonal (p. 152), and if A is real, prove that the modulus of each latent root λ is unity.

From $\bar{x}'\bar{A}' = \bar{\lambda}\bar{x}'$ and $Ax = \lambda x$, we have $\bar{x}'\bar{A}'Ax = \bar{\lambda}\lambda\bar{x}'x$. But $A = \bar{A}$, $\bar{A}'A = I$, $\bar{x}'x \neq 0$. Hence $\bar{\lambda}\lambda = 1$, which proves the result.

5. If $\bar{A}'A = I$, A is said to be unitary. Again $\bar{\lambda}\lambda = 1$.

11. **The Case of Failure for Cayley's Orthogonal Matrix** (p. 156).

From the formula $A = (I + S)(I - S)^{-1}$ we may at once deduce the correlative result $S = (A + I)^{-1}(A - I)$, by the methods of ordinary algebra, provided of course that neither $I - S$ nor $A + I$ is singular. This is so when A is any matrix whose latent roots differ from -1. Further, if A is orthogonal (real or complex), so that $A'A = I$, it follows that $S = -S'$, on simplifying the consequent expression for $A'A$; namely

$$(I - S')^{-1}(I + S')(I + S)(I - S)^{-1} = I.$$

Since there is a one to one correspondence between A and S, it follows that such an A cannot be more general than this skew symmetric $S (= -S')$. Hence Cayley's formula (37), p. 156, gives any orthogonal matrix whose latent roots differ from -1, and therefore not the most general orthogonal matrix.

If, however, A is orthogonal and $A + I$ is singular, then the formula must be modified and written

$$A = J(I + S)(I - S)^{-1}$$

as in Ex. 4, p. 157, where J is one of the 2^n matrices

$$\operatorname{diag}(\pm 1, \pm 1, \ldots, \pm 1)$$

consisting of ± 1 upon the leading diagonal and zeros elsewhere. Evidently $J = J'$, $J^2 = I$, so that $A = JB$ also is orthogonal, where $B = (I + S)(I - S)^{-1}$. Note also that $JA = B$.

To prove that *every* orthogonal matrix (real of complex) is capable of the above form it suffices to show that at least one of the 2^n matrices J allows JA to be free of the latent root -1; for then B can be taken in Cayley's form, and the result follows:

Now if $JA + I$ is singular, so is $J(A + J)$ and therefore $A + J$, since J is non-singular. If possible let all 2^n matrices $A + J$ be singular, and therefore let all their determinants $| A + J |$ vanish. Select a pair J_1, J_2 of the J which differ in sign at the first element only but are otherwise the same. Then the determinants $| A + J_1 |$ and $| A + J_2 |$ differ by twice their common first principal minor, as is seen on expanding each determinant by its top row. Hence this minor also vanishes: and this is true for all its 2^{n-1} modes, due to the further ambiguous signs. For the same reason the first principal minor of this minor vanishes, and so on until the last such minors, which are $a_{nn} \pm 1$, are reached: and certainly not both of these can vanish. Hence at least one of $| A + J |$ cannot vanish, and the theorem is proved.[1]

Example.—

Prove that every unitary matrix can be expressed by the analogous formula

$$J(I + H)\,(I - H)^{-1},$$

where H is a Hermitian matrix and $J = I$ unless H has unity for one latent root. (I owe this extension to Dr. A. C. Aitken.)

12. Matrix Representation of the Binary Symbols.

Recently it has been shown by Professor Temple how to explain the Clebsch-Aronhold symbols by means of matrices. Consider the following rectangular matrices

$$x = \left|\, x_1 \right. \qquad X_2 = \begin{bmatrix} x_1 & \cdot \\ x_2 & x_1 \\ \cdot & x_2 \end{bmatrix}, \quad X_3 = \begin{bmatrix} x_1 & \cdot & \\ & x_1 & \\ & x_2 & x_1 \\ & & x_2 \end{bmatrix}$$

where each column contains just the two elements x_1, x_2. Let continued products be formed X_2x, X_3X_2x, &c., all of which exist, since each X_r has as many columns as its predecessor X_{r-1} has rows. It is then easy to verify that

$$[a_0, a_1]\,x = a_0x_1 + a_1x_2, \quad [a_0, a_1, a_2]\,X_2x = a_0x_1^2 + 2a_1x_1x_2 + a_2x_2^2,$$

and, in general, if $a = [a_0, a_1, \ldots, a_n]$,

$$aX_nX_{n-1} \ldots X_2x = a_0x_1^n + na_1x_1^{n-1}x_2 + \ldots + a_nx_2^n,$$

which is the binary $n - ic$ with the set a for coefficients.

---

[1] This proof is due to Professor H. F. Baker. Cf. *Principles of Geometry*, 4 (Cambridge, 1940), 252.

This continued product of rectangular matrices, which has rather unexpectedly produced a scalar result, can be associated with a certain matrix, written in partitioned form as

$$B_x = \begin{array}{ccc} & aX_n & \\ & & X_{n-1} \\ & & \end{array}$$

consisting of zeros except for these rectangular sub-matrices which are arranged on the first over diagonal. For example, if $n = 3$,

$$B_x = \begin{array}{c|ccc|cc|c} \cdot & a_0x_1 + a_1x_2, & a_1x_1 + a_2x_2, & a_2x_1 + a_3x_2, & \cdot & \cdot & \cdot \\ \hline & \cdot & \cdot & \cdot & x_1 & \cdot & \cdot \\ & & \cdot & \cdot & x_2 & x_1 & \cdot \\ & & & & & & \\ & & & & & & x_1 \\ & & & & & \cdot & x_2 \\ & & & & & & \cdot \end{array}$$

Such a matrix has $\frac{1}{2}n(n + 1) + 1$ rows and columns.

Since B is linear in x_1 and x_2 it can also be written

$$B_x = B_1x_1 + B_2x_2,$$

where B_1 and B_2 are constant matrices involving the a_i, and of the same size as B. Now let

$$B_0 =$$

be a matrix of the same size with the single unit at the top right-hand corner as its sole non-zero element. By raising B_x to its nth power it is straightforward to verify that

$$B_x{}^n = \begin{array}{c} \phi \\ \\ \end{array} = \phi B_0,$$

where $\phi = aX_nX_{n-1} \ldots X_1x$, which is the binary n–ic.

It is also easy to verify that B_1, B_2 commute, and that any term such as $B_1{}^r B_2{}^s$ is equal to $a_s B_0$ when $r + s = n$, and vanishes when $r + s > n$. Thus the *matrices* B_1, B_2 behave like the symbols, say α_1, α_2, of α_x^n a binary n–*ic*; with this difference, that a combination of exactly n matrices B_1, B_2 corresponds to a matrix with a single element a_s in the top right-hand corner, instead of to the element a_s regarded as a scalar quantity.

Example.—

If $n = 2$, $a = [a_0, a_1, a_2]$, $X = X_2$, then

$$B_x = \begin{bmatrix} \cdot & aX & \cdot \\ \cdot & \cdot & x \\ \cdot & \cdot & \cdot \end{bmatrix}, \quad B_x{}^2 = \begin{bmatrix} \cdot & \cdot & aXx \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} = \begin{bmatrix} \cdot & \cdot & 1 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \end{bmatrix} \phi, \quad B_x{}^3 = 0,$$

where $\varphi = aXx = a_0 x_1{}^2 + 2a_1 x_1 x_2 + a_2 x_2{}^2$.

To polarize, we merely replace one of the x or X_i by a corresponding y or Y_i. Thus

$$aXy = ax_1 y_1 + a_1 x_1 y_2 + a_1 x_2 y_1 + a_2 x_2 y_2 = aYx.$$

In general $X_i Y_{i-1} = Y_i X_{i-1}$. Similarly for further polarization.

These matrix symbols B_0, B_1, B_2 act as a homogeneous variety of the Clebsch-Aronhold symbols, replacing 1, α_1, α_2 by B_0, B_1, B_2, and the relation $\phi = \alpha_x^n$ by $\phi B_0 = (B_1 x_1 + B_2 x_2)^n$. But they go further, for they enable us to link the symbolic theory of invariants with the abstract theory of linear associative algebra. The symbols B_0, B_1, B_2 which belong to a given binary n–*ic* are said to belong to *an algebra B*. This B, which happens to be commutative as well as associative, consists of the elements B_i each of which can be defined as a polynomial in B_0, B_1 and B_2, together with a principal element I. For our purpose we can take this I to be the unit matrix of the same size as the previous matrices. Any element B_i, being a polynomial in B_0, B_1, B_2, then reduces to a set of terms containing, at most, I, B_0 and terms of degree less than n in B_1, B_2. Those of degree n reduce to B_0, and those of higher degree vanish, as we have seen Thus every element is a linear combination $\Sigma y_i C_i$ of a *finite* number of elements C_i which are in fact

$$I, B_0, B_1, B_2, \ldots, B_1^r B_2^s, \ldots, 1 < r + s < n,$$

where the coefficients y_i are ordinary complex numbers. This set C constitutes a *basis* of this algebra, while these polynomials form a *ring*. (Cf. p. 13; and p. 235, where the basis of still more general sets of elements is considered.)

To represent more than one binary form, and to represent products of degree higher than one in the coefficients, it is necessary to introduce still larger matrices which are called *direct products* of the factor matrices representing these binary forms. Ternary and higher categories can be treated in similar fashion, but the constituent sub-matrices X_i are now more complicated.[1]

---

[1] Cf. G. Temple, "Representation of Algebraic and Differential Forms", *Journal London Math. Soc.*, **12** (1937), 114–120. D. E. Rutherford, *Proc. Roy. Soc. Edinburgh*, **62** (1943), 25–27, who introduces skew commutative complex symbols.

CHAPTER XXIII

STANDARD FORMS

1. Compound Polarization.

The theorem of p. 252 expresses any form f in p cogredient sets of n variables as a form in certain compound variables modified only by polar operators,

$$= \left(\rho_i \frac{\partial}{\partial \pi_i}\right), \tag{1}$$

where each of ρ_i and π_i is an ith compound of the original cogredient variables. The same general result may be attained, but with greater flexibility, if we replace this by the alternative operator, of the ith order,

$$\Delta' = \left(\xi\eta \ldots \,\middle|\, \frac{\partial}{\partial x}\frac{\partial}{\partial y} \ldots\right), \quad . \quad . \quad (2)$$

which has already been discussed in Chap. VII. Here $\xi\eta \ldots = \rho_i$ and $xy \ldots = \pi_i$, so that these compound variables are now resolved into their ultimate components, each consisting of i cogredient points. We may suppose that operators such as Δ' are introduced for each value of i, and that polarization of a more general type is considered as a result of replacing each such Δ by Δ'.

For example, if $f = (ab \mid xy) c_x$, then $\left(\rho_2 \frac{\partial}{\partial \pi_2}\right)$ would yield $\Delta f = (ab \mid \rho_2) c_x$, where xy in f has first been replaced by its alternative $\pi_2 = xy$, and then Δ has operated upon this first factor only, leaving c_x undisturbed. But

$$\Delta' f = \left(\xi\eta \,\middle|\, \frac{\partial}{\partial x}\frac{\partial}{\partial y}\right)(ab \mid xy) c_x$$

$$= 2(ab \mid \xi\eta) c_x + (ab \mid x\eta) c_\xi.$$

(Cf. p. 120.) The first term is a numerical multiple of Δf, whereas in the second term the variables ξ, η which appear in the operator Δ' are no longer explicitly but implicitly convolved. By a fundamental identity

$$(ab \mid x\dot{\eta})\, c_{\dot{\xi}} = (ab \mid \xi\eta)\, c_x - (abc \mid \xi\eta x).$$

Hence $$\Delta' f = 3(ab \mid \xi\eta)\, c_x - (abc \mid \xi\eta x),$$

where the first term is $3\Delta f$ and the second is a higher (here a third) compound. Now this argument is general and proceeds as on p. 121: for when Δ' contains i variables $x, y, \ldots$ and i variables $\xi, \eta, \ldots$, then $\Delta' f$ is a sum of terms T_k, where T_k is obtained from f by substituting ξ for one x, η for one $y, \ldots$, and then permuting these $\xi, \eta, \ldots$ determinantally in their $i!$ ways. The different possible choices of $x, y, \ldots$ in f necessitate the different terms T_k. If r of the $\xi, \eta, \ldots$, appear in one factor of a term, s in another, and so on, then $r!\, s! \ldots$ terms are identical; and we group them together.

Thus, in the above example, ξ replaces x in two ways, and η replaces y in one way, giving two terms. The determinantal permutation of $\xi\eta$ in the first term then yields the coefficient 2, since $(ab \mid \xi\eta) - (ab \mid \eta\xi) = 2(ab \mid \xi\eta)$.

Moreover, any term T_k, where $\xi, \eta, \ldots$ are only implicitly convolved, can be replaced by a term where they are explicitly convolved, together with terms with more than i such variables convolved. This follows from the fundamental identities involving two or more factors (p. 48 and (22) p. 344). In the case under view at least one term, T_1 say, has $\xi\eta \ldots$ explicitly convolved since we are considering a form f to which the original operator Δ applies. The requisite fundamental identity can then be written $T_k = T_1 + \Sigma T'$, where all the terms T' are of higher convolution. On equating T_k in this way to each separate term T of the series Δf, summing the results and then summing for k we obtain the further result

$$\Delta' f = \lambda \Delta f + f', \qquad \ldots \ldots \quad (3)$$

where λ is a positive coefficient and f' is a form with higher convolution. Hence each polar operator (1) can be replaced systematically by (2), which in turn can be expressed as an aggregate of simple polar operators $\left(\xi \,\middle|\, \frac{\partial}{\partial x}\right)$. Thus any form f is expressible as a series of terms $\Delta'' F$, where

$$F = \Pi\{xy \ldots st\}\, \Pi\{xy \ldots s\} \ldots \Pi\{xy\}\, \Pi\{x\} \quad . \quad (4)$$

and Δ'' is an aggregate of operators (2).

EXAMPLES

1. Show that $\left(\xi\eta \,\middle|\, \frac{\partial}{\partial x}\frac{\partial}{\partial y}\right)(ab \mid xy)^2 = 6(ab \mid xy)(ab \mid \xi\eta)$.

Here there are four terms, due to two choices of x and also of y. These terms yield the four distributions

$$\xi\eta,\ xy; \qquad \xi y,\ x\eta; \qquad x\eta,\ \xi y; \qquad xy,\ \xi\eta.$$

Determinantal permutation of ξ, η produces the term $(ab \mid \xi\eta)\,(ab \mid xy)$ in each case, but with the coefficients 2, 1, 1, 2, total 6. (Contrast the example on p. 249.)

2. $\left(\xi\eta \,\middle|\, \frac{\partial}{\partial x}\frac{\partial}{\partial y}\right)(ab \mid xy)\,(ac \mid xy) = 3(ab \mid \xi\eta)\,(ac \mid xy)$
$+ 3(ab \mid xy)\,(ab \mid \xi\eta) - (abc \mid \xi\eta x)\,(a \mid y) - (abc \mid \xi\eta y)\,(a \mid x)$.

2. Young's Standard Forms and Tableaux.

The forms F to which the general forms f have been referred are characterized by having their variables $x, y, \ldots$ in standard order. Consider such a form

$$F = (abc \mid xyz)\,(de \mid xy)\,(gh \mid xy)\,k_x \quad . \quad . \quad . \quad (5)$$

where the letters $a, b, \ldots, k$ are the symbols belonging to the original form f, but deranged. To express f we may need several such terms F (differing among themselves by the arrangement of the letters $a, \ldots, k$), together with polars of such terms. All this is comprised in the theorem of Clebsch. The next step is to find out how many of such terms F are linearly independent. This was first done by Young, who discovered a way of standardizing both the coefficients a and the variables x. By so doing he rendered the theory very precise, and brought it into relation with the theory of groups as developed by Frobenius.

Let the above form be written

$$F = F_{431} = \left(\begin{matrix} a & d & g \\ b & e & h \end{matrix}\right. \quad . \quad . \quad (6)$$

where the symbols are now arranged in separate columns instead of in the separate factors. Thus the second column *de* means that the second factor of F contains *de*, and therefore necessarily *xy* also, owing to the standard way in which the variables always appear in F. If the letters of the *j*th column are $\{p_1, p_2, \ldots\}$ in this order, then the *j*th factor of the form F must be $(p_1 p_2 \ldots | xy \ldots)$. The suffixes 431 indicate the lengths of the rows in the diagram, which is called a *tableau*. Obviously there is no lack of generality in arranging the columns in descending order of length, each column starting from the same top row. This automatically puts the rows also into descending order of length. The partitions $4+3+1$ and $3+2+2+1$ of 8, which correspond to the lengths of the rows and of the columns respectively, are called conjugate partitions. There are $8!$ ways of arranging the eight letters in the tableau corresponding to the partition $\{431\}$, but owing to the nature of the original form F many of these ways belong to the same form: for instance, the six arrangements of *abc* in the first column would yield $\pm F$ thrice each: or again, *de* may be interchanged with *gh* without altering F: but *de* interchanged with *ab* would alter F. This suggests the inquiry, how many of these $8!$ forms differ, and how many of them are linearly independent?

Let F_{431} (and generally $F_{\alpha_1 \alpha_2 \ldots \alpha_h}$) be called a *standard tableau* when the letters appear in a fixed order upon each row and in each column. The form F is then called a *standard form*. This means that initially some conventional order, say the alphabetical order, is adopted for the letters $a, \ldots, k$ and that this is to be regarded as their standard order throughout. This concept of standard order, which is of capital importance, was introduced into the theory by Young. The letters occurring in the tableau may contain repetitions, but not within the same column otherwise F vanishes off-hand.

Example.—Show that

$$\begin{pmatrix} b & a \\ c & \end{pmatrix} = \begin{pmatrix} a & b \\ c & \end{pmatrix} - \begin{pmatrix} a & c \\ b & \end{pmatrix}.$$

This expresses a non-standard form in terms of two standard forms, based on the alphabetical order a, b, c.

$[(bc \mid xy)\, a_x = (ac \mid xy)\, b_x - (ab \mid xy)\, c_x.$

3. Young's Standard Theorem.

Every Clebsch form F *can be uniquely expressed as a linear combination of standard forms which in turn are linearly independent.*

Proof. Let the letters in their standard order be denoted by 0, 1, 2, . . . respectively. If after arranging the elements down each column in ascending order the tableau is not yet standard, there must be a first element which is out of order in the tableau read by columns from left to right. For example, in

$$F = \left.\begin{array}{cccc} 0 & 0 & 1 & 0 \\ 1 & 2 & 3 & 1 \\ 4 & 3 & 4 & \\ 5 & 4 & & \end{array}\right), \qquad \begin{array}{cc} & 0 \\ & 2 \\ 4 & 3 \\ 5 & \end{array} \tag{7}$$

each column is in its proper standard order, but the 4 of the first column is the first element to be out of order. Mark with a dot this 4 and each succeeding element of its column, together with the 3, next upon its row towards the right, and all the preceding elements of its column. Since each column is in ascending order, it follows that each marked element of the first column exceeds each marked element of the next column. Now perform the operation $\dot{4}\dot{5}$, $\dot{0}\dot{2}\dot{3}$ upon these marked elements: in terms of the original form this gives an identity

$$(01\dot{4}\dot{5} \,|\, xyzt)(\dot{0}\dot{2}\dot{3}4 \,|\, xyzt) = (\ldots 45023 \ldots \,|\, xy \ldots)(\ldots \,|\, \ldots) = 0 \tag{8}$$

(p. 344) where the series on the left, consisting of $\binom{5}{2}$ terms, is necessarily equated to a set of terms with all five marked symbols explicitly convolved, and therefore with at least five of the variables x, y, . . . also convolved. Since only four such variables are distinct (F having only four rows), there must be a repetition: hence the expression vanishes.

But this identity expresses F as a sum of terms in each of which at least one marked letter of the later replaces a marked letter of the earlier column. Also if a repeated letter now occurs within a column the resulting term vanishes. In any case each new non-zero F is of the same shape, but *the sum of the integers within its first altered column is lowered.*

On repeating the whole process upon each such new F an end must be reached in a finite number of steps, since this lowering cannot go on indefinitely for a finite sum of positive integers.

Also when the end is reached and no element is out of order the elements can only be in standard order. Hence a standard form must finally be reached in every case.

Next, these standard forms are non-vanishing and linearly independent. This follows by induction upon N the number of distinct elements 0, 1, 2, ... For, in particular, let these elements represent the consecutive rows of the unit matrix, so that $a_x = 0_x = x_0$, $b_x = 1_x = x_1$, &c., where $\{x_0, x_1, \ldots\}$ are the components of x. Similarly for y, &c. A standard form then reduces to a single-term product of compound co-ordinates, say

$$F_{442} = \begin{pmatrix} 0 & 0 & 1 & 3 \\ 2 & 3 & 3 & 5 \\ 4 & 5 & & \end{pmatrix} = (xyz)_{024}\,(xyz)_{035}\,(xy)_{13}\,(xy)_{35}, \qquad (9)$$

as is easily verifiable. Delete from this tableau the final (Nth) element 5 wherever it occurs. By the nature of standard tableaux this not only leaves the rows and columns in proper diminishing lengths but leaves them still in a standard form. This form furthermore denotes $(xyz)_{024}\,(xy)_{03}\,(xy)_{13}\,x_3$, which is the complete cofactor of the product $y_5 z_5$ in the undiminished form F. Hence by induction upon N, starting with (00), that is with $x_0{}^2$, the form F is non-zero. Again, the product $y_5 z_5$, which may usefully be called the most advanced product occurring in the terms of F, will occur in any other such F, provided only that the final element 5 occurs in exactly the same places in the other's tableau. A linear relation between any F's would then imply a linear relation between the cofactors of this most advanced product, that is between standard forms of $N - 1$ distinct elements. By induction again this is impossible since, in the case of one element, but one such form, namely (00 . . .), exists and it is non-zero. This proves the theorem.

4. Enumeration of Standard Forms.

When all the elements of a tableau are distinct the number of different standard forms belonging to the same partition $\{a_1\, a_2 \ldots a_h\}$ of N is usually denoted by

$$f_{a_1 a_2 \ldots a_h}, \text{ with } a_1 \geqslant a_2 \geqslant \ldots \geqslant a_h.$$

Young found this number to be

$$N\,!\,\Pi\,(a_r - a_s - r + s)\,/\,\prod^{h}\,(a_r + h - r)\,! \quad . \quad (10)$$

where r and s run through all the $\binom{h}{2}$ pairs of values chosen from $1, 2, \ldots, h$. This follows from the fundamental difference equation satisfied by the number, namely

$$f_{\alpha_1 \alpha_2 \ldots \alpha_h} = f_{\alpha_1 - 1, \alpha_2 \ldots \alpha_h} + \ldots + f_{\alpha_1 \alpha_2 \ldots \alpha_h - 1}, \quad . \quad (11)$$

where unity is subtracted in turn from each of the h suffixes, giving h terms on the right-hand side. If, as a result, the monotonic decrease among the suffixes of a term is disturbed, the corresponding f is defined to be zero.

For instance $f_{542} = f_{442} + f_{532} + f_{541}$, but $f_{322} = f_{222} + f_{321}$ only since $f_{312} = 0$.

This difference relation is the immediate consequence of the standard construction of a tableau: for if, from all the distinct standard tableaux consisting of N different elements arranged according to the same partition $\{\alpha_1 \ldots \alpha_h\}$ of N, we delete the final (or Nth) element, we obtain all the possible tableaux enumerated by the terms on the right of (11), and no others.

This number is also given by the expression

$$f \equiv N! \left| \frac{1}{(\alpha_i - i + j)!} \right|, \quad . \quad . \quad . \quad . \quad (12)$$

where $N!$ is multiplied by an h-rowed determinant whose typical ijth element is indicated. If $\alpha_i < i - j$ the corresponding element of the determinant is zero; if $\alpha_i = i - j$ it is unity.

This is proved as follows: expand the determinant into its $h!$ terms each of which is then a multinomial coefficient $\binom{N}{p, q, \ldots, r}$ with h parts whose sum is N. By the fundamental property of such coefficients

$$\binom{N}{p, q, \ldots, r} = \binom{N-1}{p-1, q, \ldots, r} + \ldots + \binom{N-1}{p, q, \ldots, r-1},$$

whence by resumming for $p, q, \ldots, r$ the difference relation (11) is satisfied by the expressions (12). The proof is completed by induction.

Examples.—

1.

$$f_{322} = 7! \begin{vmatrix} \frac{1}{3!} & \frac{1}{4!} & \frac{1}{5!} \\[1em] \frac{1}{1!} & \frac{1}{2!} & \frac{1}{3!} \\[1em] 1 & \frac{1}{1!} & \frac{1}{2!} \end{vmatrix} = 21.$$

The reader will find it instructive to work out why the determinantal form f_{312} vanishes. In practice, to write down such determinants, first fill in the main diagonal elements, which tally with the suffixes of f; then proceed by unit steps along each row forwards and backwards, placing zeros if necessary to precede $1/0!$.

2. Prove Young's formula (10) by using Dodgson's method for evaluating the determinant.

The process is interesting and straightforward. It is well to experiment first for the cases when $h = 2$ and $h = 3$.

5. Double Standard Forms.

Another approach to this theory of standard forms is to dispense with polarization by introducing a double tableau. For brevity of exposition the trilinear form $a_x b_y c_z$ in three cogredient sets of n variables will be considered.

Let the symbols a, b, c be deranged in any manner; likewise the variables. We shall then obtain six different forms including the original, each repeated six times. Suppose that a double tableau notation be introduced and written

$$a_x b_y c_z = abc \mid xyz. \quad . \quad . \quad . \quad . \quad . \qquad (13)$$

Then $bca \mid yxz$, for instance, would denote one of these remaining five forms. What, it may be asked, is the effect of bringing these symbols bca back into standard order as before, and doing the same for the variables? The answer is that this can be done in every case provided that all possible tableaux, shaped out of three (or N) letters, are admitted, just as before. For three letters we may have the shapes $\times\times\times$, $\begin{matrix}\times\times \\ \times\end{matrix}$, $\begin{matrix}\times \\ \times \\ \times\end{matrix}$; and it is found that exactly six forms, in double standard order, are necessary:

$$abc \mid xyz; \quad \begin{matrix} \left.\begin{matrix} ab \\ c \end{matrix}\right| \begin{matrix} xy \\ z \end{matrix}, & \left.\begin{matrix} ac \\ b \end{matrix}\right| \begin{matrix} xy \\ z \end{matrix}, \\ \left.\begin{matrix} ab \\ c \end{matrix}\right| \begin{matrix} xz \\ y \end{matrix}, & \left.\begin{matrix} ac \\ b \end{matrix}\right| \begin{matrix} xz \\ y \end{matrix}, \end{matrix}; \quad \left.\begin{matrix} a \\ b \\ c \end{matrix}\right| \qquad (14)$$

Here

$$\begin{matrix} ab & xy \\ c & z \end{matrix} = (ac \mid xz)\, b_y, \quad \text{and} \quad \begin{matrix} a & \\ b & y \end{matrix} = (abc \mid xyz).$$

In this notation the single tableau, hitherto used, such as $\binom{ab}{c}$, would appear as $\begin{matrix} ab \mid xx \\ y \end{matrix}$. The proof of these results is a corollary of Young's Standard Theorem. The same systematic reduction is used upon both the symbols and the variables: the only difference is that each fundamental identity will now introduce terms which, so to speak, deepen the tableau, owing to the presence of unrepeated symbols and variables. Thus, to reduce $abc \mid xzy$ by the rule of the theorem, we start on z the first element which is out of order, and permute z, y determinantally, getting

$$abc \mid xzy = abc \mid xyz + \begin{matrix} zx \\ y \end{matrix}.$$

This introduces a deepened form, where both b and z are out of order. Two separate operations are needed, and they replace this form by the four standard forms of its own shape together with a still further deepened form $\begin{array}{c|c} a & x \\ b & y \\ c & z \end{array}$. Since these operations upon one wing, left or right, of a form either deepen the form or else leave the other wing undisturbed, they can be performed systematically, and the result is proved.

This suggests the remarkable formula

$$\Sigma f_\alpha^2 = \Sigma f^2_{\alpha_1\, \alpha_2 \ldots\, \alpha_h} = N! \tag{15}$$

or in the present case, with $N = 3$,

$$f_3^2 + f_{21}^2 + f_{111}^2 = 1 + 4 + 1 = 3!.$$

For the double standard forms can be grouped into classes as shown, one class for each shape of a tableau, and members of each class can be arranged in a square array by columns according to the f_α different standard orders of $a, b, \ldots$, and by rows according to those of $x, y, \ldots$; in all f_α^2 members. The formula, which is due to Frobenius, can then be proved on first establishing the linear independence of all these double standard forms. This follows by showing that in any assumed linear identity between the forms, any one of the forms may be left non-zero at pleasure while the rest are caused to vanish. Thus, $\begin{array}{c|c} ab & xy \\ c & z \end{array}$ being chosen, put $a = b$, $x = y$. Then all the forms to its right, or in its own class, in the arrangement (14) vanish. Operate on it with $\dot{a}$, $\dot{c}$

and all the forms (here one form only) to its left vanish. But we already know (p. 358) that the form is non-zero. Hence it is linearly independent of the rest.

6. Polarized Double Standard Forms.

Let the form $\begin{array}{cc} ab & xz \\ & y \end{array} + \begin{array}{c} ab \\ c \end{array}$, obtained by symmetrizing the row (or rows) of a double standard form as regards its variables, be called a polarized double standard form, and be written

$$\begin{array}{cc} ab & xz \\ c & y \end{array}.$$

Then
$$\begin{array}{cc} ab & z \\ c & \end{array} = \left(z\frac{\partial}{\partial x}\right)\begin{array}{cc} ab & \\ c & y \end{array} = \left(z\frac{\partial}{\partial x}\right)\binom{ab}{c}$$

in terms of the original notation for standard forms. This suggests the connexion between the present theory and the earlier work of Clebsch, Gordan and Capelli. By means of standard forms Young rendered the Gordan-Capelli series (which we have stated hitherto only in general terms) unique and precise. For the form in three sets of variable the series now runs as follows:

$$6\,a_x b_y c_z = abc\,\lfloor\underline{xyz} + 2\left\{\begin{array}{cc} ab & xy \\ c & z \end{array} + \begin{array}{c} ac \\ b \end{array}\right. \qquad + b$$

This may easily be verified. On the right there are three classes (or shapes) of term. The first class has a single term, the second class has two terms, wherein both sets *abc* and *xyz* are not only in standard form but strictly match each other, *a* with *x*, *b* with *y*, and *c* with *z*, in each term. The third class has one term. Each class has in fact f_α terms, where α as usual denotes the shape of the class. Finally the numerical coefficients of the three classes are 1, 2, 1, which are once again the values of f_α. On counting the number of forms on each side of this Gordan-Capelli series (which is the same thing as equating the 3×3 matrices $[abc]$ and $[xyz]$ to the rows and columns of the unit matrix), we revert to Frobenius' formula, here $6 = 1 + 4 + 1$, and in general $N! = \Sigma f_\alpha^2$. The Gordan-Capelli series for the case of N distinct factors a_x, b_y, . . . of the form $\phi = a_x b_y \dots d_t$ can be written uniquely as

$$N!\,\phi = \sum_{\alpha} f_\alpha \{ \sum_{\beta} [A_\beta \,\underline{|X_\beta} + \sum_{\gamma} \pm A_\gamma \,\underline{|X_\beta}] \},$$

where A_β and X_β denote corresponding standard forms within the shape α, the outer summation running through the number of different partitions α of N, the first inner summation running through the number f_α of standard forms of this same shape, while the summation for γ runs through a certain definite selection of the same standard forms each accompanied by an appropriate sign. This γ summation is absent in the instance quoted above, and makes its appearance first when $N = 5$. It is beyond our present scope to give a formal proof of this last result, the source of which—but in another notation, that of Quantitative Substitutional Analysis—may be found in Young's Third Memoir.[1] The conception of standard forms has opened up a wide field of inquiry, in which considerable new developments are still being worked out.[2] These range from the systematic investigations of D. E. Littlewood, who has carried the classical theory of generating functions a stage further by linking it with Schur's theory of associated matrices, to the geometrical applications of W. V. D. Hodge, who has solved hitherto intractable problems in enumerative geometry.

[1] *Proc. London Math. Soc.*, (2) **28** (1928), 255–292 (p. 266).
The critical step introducing what has here been called Young's Standard Theorem occurs on p. 259 of this Memoir. Cf. also Young, *Journal London Math. Soc.*, **3** (1928), 14–19.

[2] D. E. Littlewood, *Jour. London Math. Soc.*, **11** (1936), 49. *Proc. London Math. Soc.*, (2) **40** (1936), 49. "Invariant Theory, Tensors and Group Characters", *Phil. Trans.* (A) **239** (1944), 305–365, and 387–417. *Proc. Cambridge Phil. Soc.*, **38** (1942), 394–396, and **39** (1943), 197–199.
H. Weyl, *The Classical Groups* (Princeton, 1939).
H. W. Turnbull, "The Invariant Theory of a General Bilinear Form", *Proc. London Math. Soc.*, (2) **33** (1932), 1–21, and *The Theory of Critical Forms* (1942).
W. V. D. Hodge, *Proc. Cambridge Phil. Soc.*, **39** (1943), 22–30.
D. E. Rutherford, "On the Relations between the Numbers of Standard Tableaux", *Proc. Edinburgh Math. Soc.*, (2) **7** (1942), 51–54, and "On Substitutional Equations", *Proc. Roy. Soc. Edinburgh*, **62** (1944), 117–126.

INDEX

Absolute, 320.
Absolute invariants, 206–7, 275.
Addition of matrices, 34, 70.
Adjacent terms, 15, 251
Adjugate, 67, 104.
  compound, 87.
  Jacobi's theorem on the, 79.
Adjunction theorem, 323.
  for affine group, 310.
  for orthogonal group, 316.
Affine group, 162, 309.
  invariants, 227, 310.
  symbolic expression for, 310.
  transformation, 309.
  with fixed point, 162, 315.
Aitken, 29, 108.
Algebra, definition of, 281, 351.
Algebraic complement, 26.
Algebraically complete system, 231, 234.
Alternant, 28.
Alternation. See *Determinantal permutation.*
Analysis, 281, 326.
Anharmonic ratio. See *Cross ratio.*
Annihilators, 227.
Apolarity, 262–9, 278.
  defined, 264.
Argand diagram. See *Gauss plane.*
Arithmetic, 281.
Aronhold, 141, 174, 179.
Aronhold operator, 140, 180, 207, 260, 302.
Associative law, 57, 61.
Ausdehnungslehre, 58.
Axial co-ordinates, 87.

Basis theorem of Hilbert, 235.
Battaglini complex, 289.
Bazin, 55, 56, 108, 225.
Bell, 177, 181.
Bernoulli, 13.
Bezout, 46.
Bideterminant, 82, 342–6.
Bilinear forms, 102, 198, 294.
  binary and higher, 286.
  geometrical interpretation of, 283–7.
  latent in transformation, 159.
Binary field, 10, 41, 50, 59, 128–46, 151, 173, 177–81, 191, 213, 215–33, 244–7, 263, 266, 269, 283–6.
  double, 243.
Binet-Cauchy theorem, 79–81, 341.
Bôcher, 160, 296, 304.
Boole, 128, 139.
Bromwich, 304.
Burchardt, 325.

Canonical forms, 265.
  binary, 219, 244, 246.
  of general quadric, 300, 302.
Canonical matrix, 296.
Capelli, 112, 116, 253.
  operator, 112, 256.
Cartesian co-ordinates, 3, 5, 128–32, 155, 163, 178, 280, 283, 310, 325, 328.
Cauchy, 55, 56, 67, 87, 108, 165.
Cayley, 2, 4, 59, 132, 133, 155, 156, 174, 227, 234, 296, 348.
  Hamilton theorem, 99, 296, 331.
  operator, 113–5, 122, 123, 188, 203, 211.
Characteristic of determinants, 55.
  equation, 98, 107, 292, 347.
Chinese, 6.
Ciamberlini, 307.
Circle, 243.
Class $C+$, $C-$, 15.
Clebsch, 140, 172, 174, 179, 210, 247, 248, 255, 258, 260.
  theorem of, 248.
  transference principle, 287, 299, 316.
Cogredient, transformation defined, 149.
  sets of variables, forms with, 199, 270.
Collineation, 291, 295.
Combinants, 242.
Commutant, 332.
Commutative law, 57, 61, 110, 257.
  matrices, 71.
Complement, algebraic, 26.
Complete systems, 243.
  algebraically, 231.
  of binary cubic, 244.
  of binary quartic, 245.
  of binary various, 246.
  of general quadric, 297.
  two general quadric, 304–6.
Complex, linear, 212, 329.
  quadratic, 258.
  quaternary line, 329.
Complex variable, 243.
Compound co-ordinates, 85, 86, 209, 249.
  determinants, 49, 87.
  inner product, 81, 83, 209.
  transformation, 163, 165.
Cone, 299.
Conformable matrices, 34, 336.
Conic as binary form, 284.
  as ternary, 178, 286, 289, 298, 302.
  polar, 178, 286.
Conjugate matrices, 5.
  points, 287, 289.
  primes, 287.
Continued fractional matrix, 331.
Contracted functional notation, 20.
Contragredience, defined, 149.
  and correlation, 295.
  fundamental property of, 149.
  of point and prime, 150, 151.
Contravariant, 206, 329.
Convolution, implicit, explicit, 46, 225, 253.
  and resolution, 207.
  double, 193.
Co-ordinates. See *Cartesian, Compound.*
  homogeneous, 5, 86, 151, 212, 265, 283–308, 315, 325.

Correlation, 295.
Correspondence, 280, 281, 291.
Corresponding matrices, theorem of, 79, 116.
Covariants of binary forms, 143.
  as invariants of linear forms, 145, 207.
  defined, 144.
  in the relativity theory, 329.
  of degree two, 230, 274.
  of general forms, 206.
Cramer, 13, 31.
Cross ratio, 283.
Cubic, binary, 244, 266.
  syzygy, 244.
  twisted, 286.
Cullis, 296.
Currency, defined, 37.

Degree defined, 134, 172.
Derangement, 13.
Desargues, 332.
Determinant. See *Adjugate*, *Alternant*, *Compound*.
  bordered, 101, 299.
  characteristic, 55.
  definition of, 17.
  differentiated, 110, 123.
  duality of, 51, 89, 92.
  expansion of, 19, 98.
  extensional, 42–9, 287–9.
  functional. See *Jacobian*.
  irresoluble, 33.
  logarithm of, 124.
  multiplication of, 65.
  notation of, 1, 27, 37.
  reciprocal, 66, 89, 103.
  skew symmetric, 105, 348.
  symmetric, 104.
Determinantal permutation, 27, 38, 43–5, 48–51, 210.
  as a differential operation, 121.
Diagonal matrix, 101.
Dickson, [illegible], [illegible], 304.
Differential equation satisfied by determinantal series, 123.
  satisfied by normal forms, 256–8.
  forms, 327.
  invariants, 329.
Differentiation of a determinant, 123.
Dimensions of a group, *161*.
Discriminant, 129, 131, 140, 191, 192, 218.
  of conic, 287, 298, 322.
Distributive law, 57, 61.
Division law, 57, 64.
Dodgson's method, 340, 360.
Double binary forms, 138, 243.
  convolution for quadrics, 193.
  suffix notation, 63.
Duality, principle of, 262, 282, 284, 298, 301.
  and determinants, 51, 92.
  formal, 54.

Elimination leads to invariants, 245, 274.
Elliott, 232, 246.
Elliptic geometry, 315.
Equations, linear, solution of, 10, 75, 84, 195.
Equivalence problem, 277.
Equivalent forms, 277.
  symbols, 179, 191, 201.
Euclidean geometry, 315, 316.
Euler, 111, 157, 256.
Extensionals, 42, 49, 287.
Extrinsic terms, 116.

Factor. See *Symbols*.
Ferrar, 55.
Field. See *Binary*, *Ternary*.
  currency, 37, 54.
  number, 9.
Finiteness theorem, 233–40.
Fore and after factors, 64.
Form, 30, 133, 168, 176.
Formal dual, 54.
Forsyth, 233.
Frame of reference, 293.
Franke, 108.
Frobenius, 64, 128, 258, 330, 347, 361.
Fundamental theorem, first, 182.
  second, 214, 225, 314.
  affine, 312.
  for general forms, 190, 203, 208.
  for linear forms, 182, 187.
  Hermitian, 322.

Gauss, 128.
Gauss plane, 285.
General forms, 181.
Generator of quadric, 299.
Geometry, 280–96. See *Cartesian*, *Co-ordinate*.
Gilham, 223, 307.
Gordan, 172, 193, 214, 234, 238, 242, 243, 247, 255, 258, 307, 314.
Gordan's theorem, 233, 261.
  proof of, 238.
Gordan-Capelli series, 253, 254, 255, 270, 272, 313, 363.
Grace, 273, 274, 278.
Grace and Young, 114, 234, 238, 242, 246, 253, 261, 284, 307.
Gradient, 134, 231.
  expressed as sum of coefficients of covariants, 270.
Gram, 270–4.
Grassmann, 58.
Ground forms, 172.
Group defined, 160.
  affine, 161, 309.
  general projective, 161.
  Hermitian, 320.
  homogeneous, 309.
  isobaric, 310.
  Lorenz, 322.
  orthogonal, 160, 322.
Group property, 283.

Hamilton, 58, 99.
Harmonic range, 283.
  invariant of quadratic, 316.
  extensional of, 287.
Hermite, 159, 320.
Hessian, 181, 222, 231, 233.
  identical vanishing of, 274.
Hilbert, 172, 234, 235, 240, 261, 314.
Hodge, 363.
Hölder, 57.
Homogeneity, 30, 280. See *Co-ordinates*.
  of invariants, 171.
Homographic ranges and transformations, 291.
Hyperbolic geometry, 315.
Hyperdeterminants, 132.

Identical transformation, 161, 270.
Identities, fundamental, 44, 334.
  binary, 213, 278.
  dual, 93, 344.
  general, 214, 278.
  Laplace, 41–56, 93.
  Sylvester, 45, 94–7.
Improper orthogonal group, 322.
Index law, 68.
Induced transformation, 135.
Inner product, 62, 82, 83, 316.
  compound, 83.
Integration of a rational function, 28.
Intrinsic terms, 116.
Invariant, defined, 138, 169.
  as elimination result, 275.
  as solution of differential equation, 230.
  equations, 271.

Invariant factors, 304.
  of binary forms, 128, 244.
  of [illegible] forms, 184, 189.
  of multiple fields, 240.
  of affine group, 309.
  of orthogonal group, 130, 316, 320.
  process, 141.
  projective, 129, 139, 169.
Inverse transformation, 135.
Inversion, 243.
Involution, 219, 284, 292.
Irreducible systems of forms, 233. See *Complete systems.*
Isobaric. See *Weight.*

Jacobi, ratio theorem of, 77, 89, 108, 339.
  lemma of, 125.
Jacobian, 124–7, 151, 231, 242, 327.
  and canonical forms, 268, 302.
  is a covariant, 143, 182.
  of a Jacobian is reducible, 223.
  of binary forms, 143, 181, 219, 221, 284.
  of two quadratics is harmonic to both, 284.
  product of two, 223, 307.
  rank of, 126.
  vanishing of, 126.
Jessop, 304.
Jordan's lemma, 279.

Kasner, 243.
Klein, 224.
Kronecker, 33.

Lagrange, 128, 229.
Laplace's development of a determinant, 22, 83, 323, 336.
Lasker, 267.
Latent points, 292, 296, 315.
  primes, 314.
  quadric, 315.
  roots, 98, 101, 292, 347.
Leading coefficient, 226.
  diagonal expansion by, 98.
Lehnen, 243.
Levi-Civita, 328.
Line co-ordinates, 85, 86, 285.
  geometry, 85.
Linear dependence, 8, 50, 73.
  equations 6, 10, 285.
  forms, 31.
  forms, invariants of, 145.
Linearity, 30.
Littlewood, 363.

Macmahon, 119.
Matrix, defined, 2.
  commutative, 71, 332.
  diagonal, 101.
  Hermitian, 348.
  Jacobian, 327.
  null, 5.
  orthogonal, 152, 153–7, 331, 347–9.
  reciprocal, 68.
  scalar, 71.
  singular, 70.
  unit, 68.
Matrix properties, 34, 59, 70.
  and quaternions, 166.
  canonical form of, 296.
  function of, 71.
  geometrical interpretation of, 291, 295.
  transformation, [illegible].
Mertens, 258.
Metrical properties, 324.
Meyer, 234, 242, 308.
Minor determinant, 21, 26.
Mixed concomitant, 206.
Modulus of transformation, 169, 311.
Muir, 29, 46, 87.
Multilinear forms, 197–201.
Multilinear invariants, 142, 261.
Multiple fields, 240.
Multiplication of matrices, 59–62.
  scalar, 61.
  fore and after, 64.

Napier, 331.
National independence, 234.
Net, 260.
Noether, 247, 258, 261.
Non-commutative, 59, 119.
Norm, 166.
Norm curve, 285.
Normal form, 255.
Null matrix, 5, 61.
  system, 295.

Operator. See *Aronhold, Capelli, Cayley.*
  differential, 227.
Order of matrix, 1, 7, 17.
  of polynomial, 30, 133.
Orthogonal. See *Group, Matrix.*
  invariants, 315.
Outer product, 183, 316.

Pappus, 333.
Parallel property, 310.
Partial fractions, 28.
Pascal, 214.
Peano, 243, 260.
  theorem of, 261, 314.
Pencil, 218, 260.
Permanent, 14, 251.
Permutation, 13.
  determinantal, 27.
Perpetuant, 246.
  double binary, 243.
Picquet, 109.
Platonic solid, 224.
Polar, adjacent terms of, 251.
  forms, 137.
  prime, 286, 294.
  reciprocation, 295.
  symbolical expression for, 177.
Polarization, 110, 353.
  an invariant process, 207, 250.
Polynomial function of matrix, 71.
Prepared system, 329.
Prime, 86, 163, 282.
Product of matrices, 61, 63, 71.
  inner, 62.
  outer, 183.
Projection, 290.
Projective property, 271, 290.
  determined by invariant equation, 272
Proper orthogonal group, 316.
Pythagoras, 325.

q-numbers, 59.
Quadratic. See *Complete system, Conic.*
  as determinant, 105.
  [illegible] in transformation, 158, 315.
  reduction, 194.
Quantic, 133.
Quartic, 245.
Quaternary variables, 262.
Quaternion, 166.

Rank of matrix, 5, 73, 75, 84.
  of quadric, 299.
Rational curve, 285.
Rationality, 10, 12, 157, 233 320.
Reciprocal matrix, 68.
Reciprocation, 16, 283, 294
Reducibility, 225.
Real, 108, 250.
Relative invariant, 129, 277.
Relativity, [illegible], 233, 326.
Resolution, 208.
Restricted transformation, 227.

Resultant, 275.
Reversal law, 68.
Riemannian geometry, 315, 329.
Rigid displacement, 131, 155.
Rodrigues, 157.
Rothe, 29.
Rutherford, 352, 363.

Saddler, 243.
Salmon, 132, 302.
Scalar instance of symbols, 175, 177.
Scalar matrix, 71.
Schur, 363.
Schwartz, 243.
Schweins' series, 339.
  theorem, 336.
Self-conjugate simplex, 299, 304.
  triangle, 299.
Seminvariant, 226.
Similar forms, 259.
  dual forms, 262.
Simplex, 84, 293, 299, 304.
Singular matrix, 70.
  point, apolarity theory of, 265.
Skew symmetry, 105.
  of determinants, 105–7.
  of matrices, 36.
  of null system, 295.
Smith, 33.
Space of $n - 1$ dimensions, 282.
Standard forms, 250 353–8.
Standard order, 19.
Straight line, geometry on, 283.
Stroh's lemma, 278.
Study, 214, 243, 258, 325.
Subgroup, 161, 309.
Submatrix, 38.
Substitution, 128.
Substitutional analysis, 119, 330.
Summary of matrix laws, 71.
  theorems on compound determinants, 108.
Sylvester, 46, 55, 56, 87, 108, 121, 132, 133, 165, 227, 234, 250.
Symbols, defined, 173, 175, 198, 349.
  contragredient, 200.
  effect of linear transformation on, 183, 201.
Symbolic factor types, 182.
Symbolic linear equation, solution of, 196.
Symmetric function of roots, 134.
  matrix, 36.
Syzygy, 224, 231.
  cubic, 244.
Syzygy, finiteness of system of, 239.
  for quadratics, 220.
  quartic, 246.

Tableau, 355.
Tangential equation, 285.
Temple's symbols, 349–52.
Tensor, 89, 90, 200, 329.
Transference principle. *See Clebsch.*
Transformations, linear, 59, 128, 168.
  defined, 147.
  form a group, 160.
  general functional, 151, 326.
  induced, 135–7, 148.
  See *Group.*
Transposed matrix, 5, 36, 71.
Transposition, 70.
Transposition properties of determinants, 38.
Transvectants, 221.
Turnbull, 46, 55, 243, 307, 363.
Types, 217.

Unit determinant, 32.
  matrix, 61, 68, 339.
Unitary matrix, 344.
Upper suffix notation, 77, 89, 200.

Vaidyanathaswamy, 243.
Valency condition, 186, 190, 191, 231.
Van der Waerden, 225.
Variables, dual, 90.
  compound, 86.
Veblen, 328.
Vector, point, 36, 84, 86, 91.
  of orthogonal group, 317.
  prime, 36, 84, 86, 91.
  properties of, 59.
Von Gall, 246.

Wakeford, 267.
Weierstrass, 105.
Weight, 134, 170, 310.
  homogeneity of, 171.
Weitzenböck, 46, 214, 225, 239, 240, 258, 261, 308, 309, 315, 325, 328, 363.
[illegible]ker, 55, 56.
Williamson, 307.

Young, 247, 258, 330, 355. See *Grace and Young.*